

BSi

PEAKSTM *Studio 5.1*

User's Manual

superior
proteomic ms/ms
software

peak detection & deconvolution
inChorus

de novo sequencing
quantification

peptide homology discovery

protein identification

data refinement
batch workflow



BIOINFORMATICS SOLUTIONS INC.

PEAKS Studio 5.1

User's Manual

© Bioinformatics Solutions Inc.
470 Weber St. N. Suite 204
Waterloo, Ontario, Canada N2L 6J2
Phone 519-885-8288 • Fax 519-885-9075

Please contact BSI for questions
or suggestions for improvement.

1. Introduction to PEAKS 5.1	6
1.1 Main Features.....	6
1.2 New Features in the PEAKS 5 Product Line	6
1.3 Workflow	7
1.4 Guidelines for Using this Manual	7
1.5 Scope.....	7
1.6 Service and Support.....	7
2. Getting Started with PEAKS 5.1	8
2.1 Package Contents	8
2.2 System Requirements	8
2.2.1 Adjusting the Amount of Memory that PEAKS Uses	8
2.3 Installation for Windows Users	9
2.4 Registering PEAKS.....	10
2.4.1 Registration Instructions (Internet Connection).....	10
2.4.2 Registration Instructions (No Internet Connection).....	11
2.4.3 Re-registration Instructions	14
2.5 Set up PEAKS Preferences	14
2.6 Set up PEAKS Configuration	14
3. Quick Walkthrough	15
3.1 Create a Sample Database.....	15
3.2 Create a Project.....	15
3.3 Perform Data Refinement	17
3.4 Run <i>de novo</i> Sequencing.....	18
3.5 Run Protein Identification	21
3.6 Run PTM Finder.....	23
3.7 Run an inChorus Search	25
3.8 Perform a SPIDER Search.....	30
4. Load Data.....	33
4.1 Data Formats.....	33
4.2 Data Conversion.....	33
4.2.1 PEAKS Project Converter	33
4.2.2 Agilent Data.....	34
4.2.3 Applied Biosystems (WIFF)Data.....	34
4.2.4 Applied Biosystems (4700/4800)Data	35
4.2.5 Bruker Data	35
4.2.6 Shimadzu Data	35
4.2.7 Thermo Data.....	36
4.2.8 Varian Data	36
4.2.9 Waters/Micromass (MassLynx) Data	36
4.3 Create a New Project	37
4.4 Add Data to an Existing Project	38
4.5 Open / Close a Project	39
4.6 Delete and Hide Results.....	39
4.7 Change the Location of Saved Projects	39
4.8 Orienting Yourself	40

4.8.1 Project View Panel.....	40
4.8.2 Properties Panel	40
4.8.3 Raw Spectrum View	41
5. Data Refinement.....	43
5.1 Run Data Refine.....	43
5.2 Data Refinement Parameters.....	44
5.2.1 Merging Scans	44
5.2.2 Precursor Charge Correction	44
5.2.3 Filtering MS/MS Scans.....	44
5.2.4 Preprocessing MS/MS Scans.....	44
5.3 Data Preprocessing Results	44
6. <i>de novo</i> Sequencing	45
6.1 Setting up Auto <i>de novo</i> Sequencing Parameters.....	45
6.1.1 Mass Options	46
6.1.2 Enzyme Options	46
6.1.3 PTM Options	46
6.1.4 General Options	46
6.2 <i>de novo</i> Sequencing Results.....	47
6.2.1 Peptide Candidates Frame	48
6.2.2 Ion Table Frame.....	50
6.2.3 Spectrum View Frame.....	50
6.2.4 Spectrum Alignment Frame.....	51
6.2.5 Survey Scan	51
6.2.6 Error Map.....	51
6.3 Manual <i>de novo</i> Sequencing.....	51
6.3.1 Manual <i>de novo</i> Graphical User Interface.....	51
6.3.2 Manual <i>de novo</i> Operations	53
7. Database Search	58
7.1 Setting up Protein Identification Parameters.....	58
7.1.1 Mass Options	59
7.1.2 Enzyme Options	59
7.1.3 General Options	59
7.1.4 PTM Options	60
7.1.5 Database Options	61
7.1.6 Advanced Options.....	61
7.2 Protein Identification Results	62
7.2.1 Peptide View	62
7.2.2 Protein View	64
8. SPIDER Search	69
8.1 Setting up SPIDER Parameters.....	69
8.1.1 Query Options	70
8.1.2 General Options	70
8.1.3 PTM Options	71
8.1.4 Filter Options	71
8.1.5 <i>de novo</i> Options	72
8.1.6 Database Options	72

8.2 SPIDER Results View.....	72
9. PTM Finder	74
9.1 Setting up PTM Finder Parameters.....	74
9.2 PTM Finder Results View.....	75
10. inChorus Meta Search.....	76
10.1 Setting up inChorus Parameters	76
10.1.1 Importing Existing Results	77
10.2 inChorus Results View	78
10.2.1 <i>de novo</i> , Peptide and Protein Views.....	78
11. Creating a High-Throughput Work Flow	79
11.1 Setting up the Work Flow Parameters.....	79
12. Filtering Your Results	81
12.1 Setting Filter Parameters	81
12.1.1 Possible Filters/ Selected Filters/ Edit Filter	82
12.1.2 Filter Options	83
12.1.3 Parameter Options.....	84
13. Complex Analysis.....	86
13.1 Creating a project for complex system	86
13.2 Integrating data analysis	86
14. Exporting Data/Reports and Printing.....	87
14.1 Export Data in .mzxml or .mgf.....	87
14.2 Export Results File.....	87
14.3 Export Peptide Results in PepXML Format	88
14.4 Export Results in Excel Format.....	88
14.5 Print Tables and Graphs for Publication	89
14.5.1 <i>de novo</i> Image Files	89
14.5.2 Protein ID Image Files.....	90
15. Advanced Configuration and Environment Preferences.....	92
15.1 PEAKS Environment Preferences.....	92
15.1.1 General Preferences.....	92
15.1.2 Instrument Preferences	94
15.1.3 Search Engine Preferences.....	96
15.1.4 Ion Editor Preferences.....	98
15.2 PEAKS Configuration	99
15.2.1 Enzyme Configuration.....	99
15.2.2 PTM Configuration	100
15.2.3 Database Configuration.....	101
15.2.4 Instrument Configuration	105
15.2.5 Parameter Configuration	107
16. PEAKS Quantification Module	108
16.1 Setting up PEAKS Q Parameters.....	108
16.2 3D View	110
16.3 iTRAQ Walkthrough.....	111
16.3.1 Setting up a Database	111
16.3.2 Creating a Project.....	112

16.3.3	Setting up the Workflow Parameters.....	113
16.3.4	ITRAQ Results	118
16.4	SILAC Walkthrough	119
16.4.1	Setting up a Database	119
16.4.2	Creating a project	120
16.4.3	Setting up the Workflow Parameters.....	121
16.4.4	SILAC Results	125
16.5	Label Free Quantification	127
16.5.1	Setting up a Database	127
16.5.2	Creating a Project	128
16.5.3	Setting up the Workflow Parameters.....	129
16.5.4	Label Free Results.....	132
17.	References	135
18.	Appendix	136
18.1	Terminology and Abbreviations Glossary	136
18.2	Toolbars	137
18.2.1	Main Window Toolbar	137
18.2.2	Project View	139
18.2.3	Main Processing Window Toolbar	139
18.3	Mass Calculator	140
18.3.1	Advanced Options.....	141
19	About Bioinformatics Solutions Inc.	142
20	PEAKS Software License	143

1. Introduction to PEAKS 5.1

1.1 Main Features

PEAKS is an innovative software system designed to derive amino acid sequences and identify proteins using tandem mass spectrometry data from all major mass spectrometry vendors. PEAKS incorporates *de novo* sequencing results into the database searching process for peptide/protein identification. It does this by generating sequence tags which are used in conjunction with fragment ion mass matching to speed up the search, remove false positive matches, and find peptides with interesting sequence variations or modifications that would prevent them from being otherwise identified. Our meta protein search tool, inChorus allows users to use multiple search engines (PEAKS, X!Tandem, OMSSA Sequest and Mascot) to expand sequence coverage and increase confidence. Another feature is SPIDER, which is a homology search tool for *de novo* sequencing. SPIDER is also able to reconstruct the correct sequence using the *de novo* sequence and a homologous peptide.

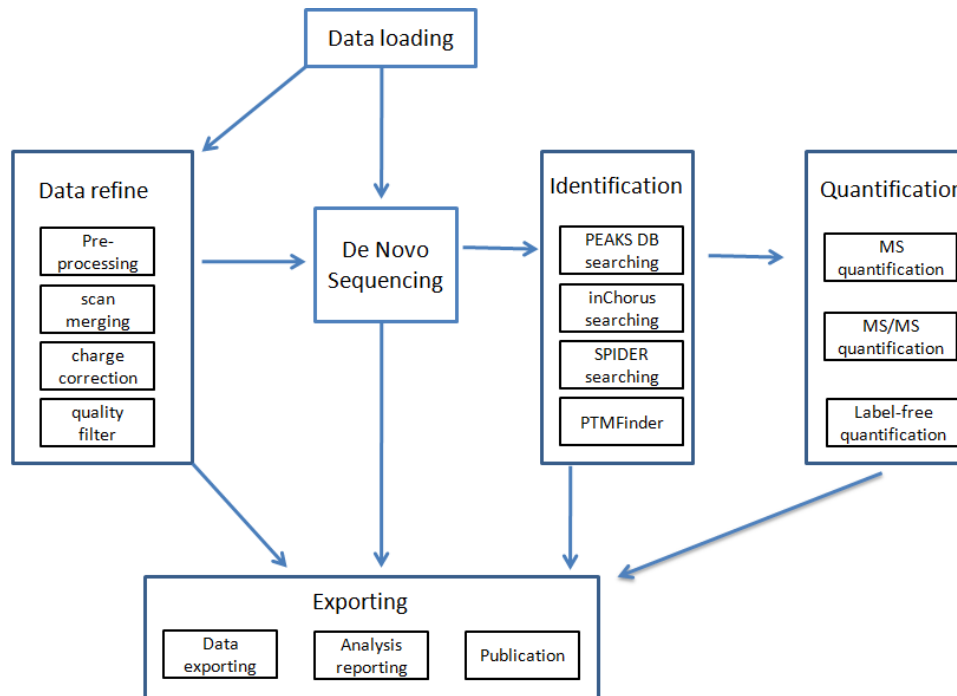
1.2 New Features in the PEAKS 5 Product Line

We have many new features in PEAKS 5.1 which will be explained throughout this manual. One main feature which required a complete redesign of the program architecture is the ability to handle very large data sets. In addition, our protein identification is more sensitive and generates less false positives.

BSI, creator of PEAKS, have also produced a quantification module that will allow users to automatically quantify proteins from experiments using both label and label-free techniques. Results generated from PEAKS Q have high accuracy and can be performed over a wide dynamic range with very large datasets. PEAKS 5.1 also has improved identification of PTMs with our innovative PTM finder.

PEAKS 5.1 uses project based data management which allows users to process simultaneous runs and easily compare/contrast samples within one project.

1.3 Workflow



1.4 Guidelines for Using this Manual

This user's manual is intended to help you get started with PEAKS 5.1. It will describe its functionalities, show how to customize PEAKS to your applications, provide a task based reference, and troubleshooting. We recommend reading the walkthrough in Chapter 3, using the sample data provided.

1.5 Scope

PEAKS users are assumed to be familiar with computer usage and the operating system environment. As such, it is beyond the scope of this manual to instruct the user on the use of windows, dialogue boxes, menus, file storage etc. Please refer to the operating system's manual or computer help books for such information. Similarly, PEAKS users are expected to be familiar with mass spectrometry, standard operating practices and data.

1.6 Service and Support

In addition to reading this manual, we recommend that you take a look at our training videos that explain the main features of PEAKS in detail (<http://www.bioinfor.com/products/peaks/support/tutorials.php>). Please send technical questions to support@bioinfor.com.

We also encourage our users to provide us with any suggestions or comments as we are always improving PEAKS to meet the needs of the scientific community (<http://www.bioinformaticssolutions.com/corporate/contactform.php>).

2. Getting Started with PEAKS 5.1

This section of the manual will guide you through the installation and configuration of PEAKS 5.1.

2.1 Package Contents

The PEAKS 5.1 package contains:

- This manual (hardcopy and/or electronic version)
- PEAKS 5.1 software

2.2 System Requirements

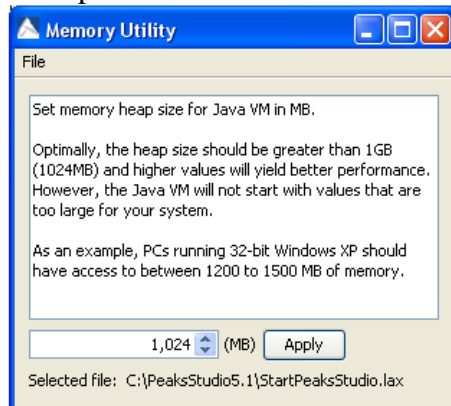
PEAKS 5.1 will run on most platforms with the following minimum requirements:

- Processor: Equivalent or superior processing power to an IntelTM Pentium 4 Processor 1.6GHz (we recommend an IntelTM CoreT2 Duo processor/AMD PhenomTM processor)
- Memory: 1 GB memory (we recommend 2GB memory)
- Operating System: PEAKS runs on Windows XP and Vista.

2.2.1 Adjusting the Amount of Memory that PEAKS Uses

The PEAKS Studio directory (e.g., C:/PEAKS Studio 5.1) contains a file called Memory Utility.exe. Click to open this file and the following window will open:

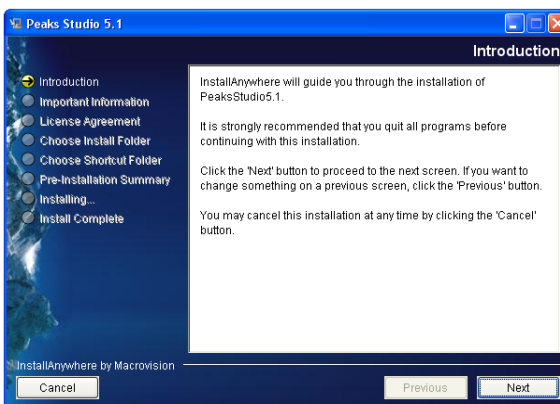
The default of 1024 MB tells the Java Virtual Memory (which runs PEAKS) to run with 1024 MB of memory. You can increase the Java VM to determine the highest optimal value for your own computer although some trial and error may be needed as Java will not start if you set the value to be too large for your system. Please do not hesitate to contact us if you have any questions about your own system.



2.3 Installation for Windows Users

Note: If you already have an older version of PEAKS installed on your system, please uninstall it before proceeding.

1. Close all programs that are currently running.
2. Insert the PEAKS 5.1 disc into the CD-ROM drive. If loading PEAKS via the download link, skip to step 4, after downloading and running the file.
3. Auto-run should automatically load the installation software. If it does not, find the CD-ROM drive and open it to access the disc. Click on the PEAKS_Studio_Installation.exe.
4. A menu screen will appear. Select the top item “PEAKS Installer.” The installation utility will begin the install. Wait while it does so. When the “PEAKS 5.1” installation dialogue appears, click the “Next” button.
5. Basic system requirements will be presented. Please note that while the minimum requirement is 1 GB of RAM, the preferred configuration is 1.5 GB of RAM. “Click Next”.
6. Read the license agreement. If you agree with it, change the radio button at the bottom to select “I accept the terms of the License Agreement” and click “Next”.
7. Choose the folder/directory in which you would like to install PEAKS. The default location is simply “C:\PeaksStudio5.1. To change this location press the “Choose...” button to browse your system and make a selection, or type a folder name in the textbox. Please avoid installing PEAKS in the “Program Files” directory as well as in any directory for which the PEAKS user will not have write-permissions. Click “Next”.
8. Choose where you would like to place icons for PEAKS 5.1. The default will put these icons in the programs section of your start menu. A common user preference is on the desktop. Click “Next”.
9. Review the choices you have made. You can click “Previous” if you would like to make any changes or click “Next” if those choices are correct.
10. PEAKS 5.1 will now install on your system. You may cancel at any time by pressing the “Cancel” button in the lower left corner.



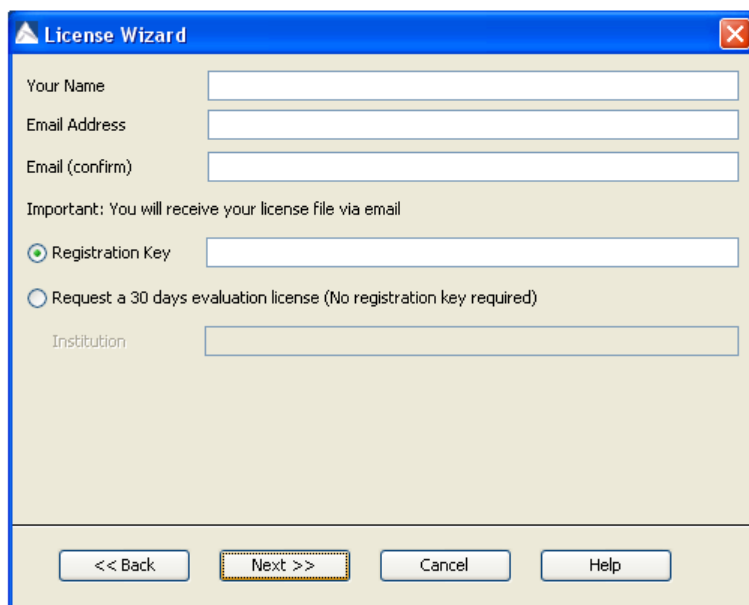
11. When the installation is complete, click “Done”. The PEAKS 5.1 menu screen should still be open. You may view movies and materials from here. To access this menu at a future date, simply insert the disc in your CD-ROM drive.

2.4 Registering PEAKS

The first time PEAKS is run, the license wizard will appear automatically:

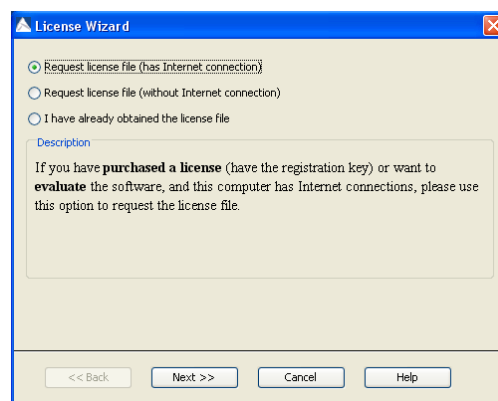
2.4.1 Registration Instructions (Internet Connection)

1. Select “Request license file (has Internet connection)” and click “Next”.
2. The following window will appear:



The screenshot shows the 'License Wizard' window. It has a title bar with a triangle icon and a close button. The main area contains several input fields: 'Your Name', 'Email Address', 'Email (confirm)', 'Registration Key', and 'Institution'. There are two radio buttons: 'Registration Key' (selected) and 'Request a 30 days evaluation license (No registration key required)'. Below the radio buttons is a text box for 'Institution'. At the bottom, there are four buttons: '<< Back', 'Next >>' (highlighted with a yellow border), 'Cancel', and 'Help'. A message above the radio buttons says: 'Important: You will receive your license file via email'.

An automated BSI service will generate the license file (license.lcs) and email it to the provided email account from the License Wizard. You can either save the attachment to a local directory or copy the content between '====>' and '<====' in the email. Click “Next”.



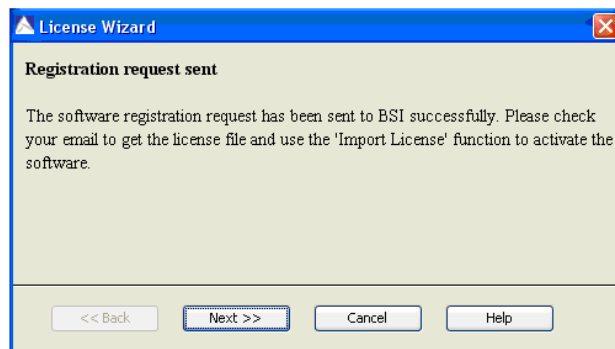
The screenshot shows the 'License Wizard' window. It has a title bar with a triangle icon and a close button. The main area contains three radio buttons: 'Request license file (has Internet connection)' (selected), 'Request license file (without Internet connection)', and 'I have already obtained the license file'. Below the radio buttons is a text box with the following text: 'If you have purchased a license (have the registration key) or want to evaluate the software, and this computer has Internet connections, please use this option to request the license file.' At the bottom, there are four buttons: '<< Back', 'Next >>', 'Cancel', and 'Help'.

If you have purchased PEAKS and have a registration key, select “Registration Key”. Enter your registration key as well as your name and email address and click “Next”.

OR

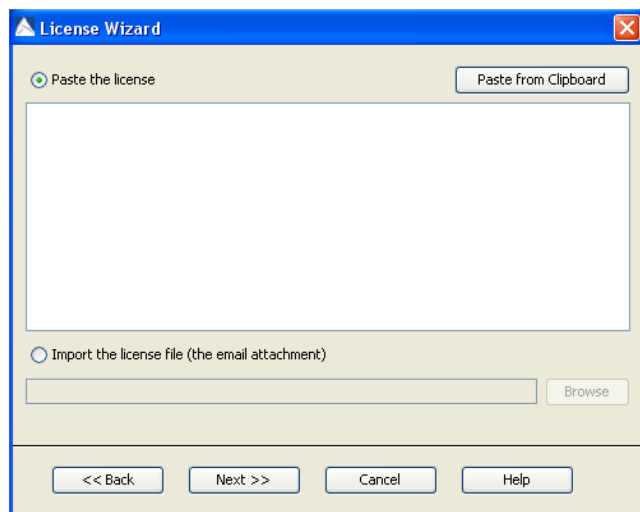
If you are trying a demo of PEAKS and do not have a registration key, select “Request a 30 days evaluation license (No registration key required)”. Enter your name, email address, as well as your institution. Click “Next”.

3. The following window will appear:



The screenshot shows the 'License Wizard' window. It has a title bar with a triangle icon and a close button. The main area contains the text: 'Registration request sent' followed by 'The software registration request has been sent to BSI successfully. Please check your email to get the license file and use the 'Import License' function to activate the software.' At the bottom, there are four buttons: '<< Back', 'Next >>' (highlighted with a yellow border), 'Cancel', and 'Help'.

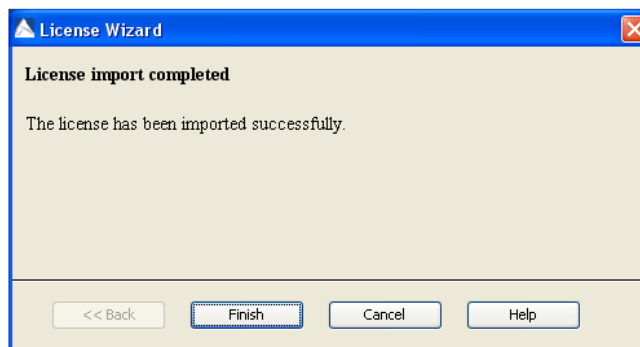
4. The following window will appear:



Select “Paste the license and click on the “Paste from Clipboard” button to paste the license information between '===>' and '<====' in the email or select “Import the license file (the email attachment) and browse to locate the license file (license.lcs). Click “Next”.

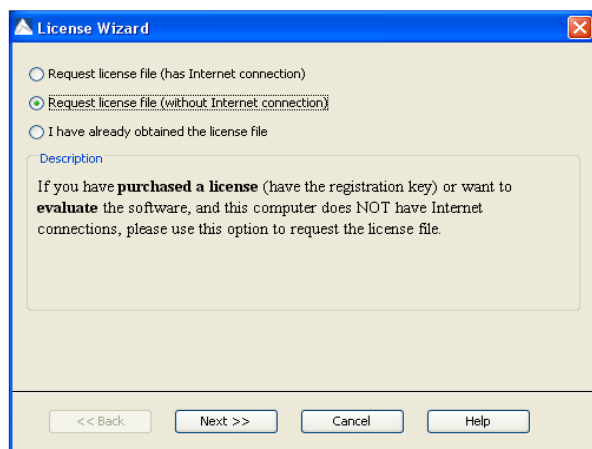
5. The following window will open:

Click “Finish” if you receive a message that the license has been imported successfully.



6. Restart PEAKS to begin using the software.

2.4.2 Registration Instructions (No Internet Connection)



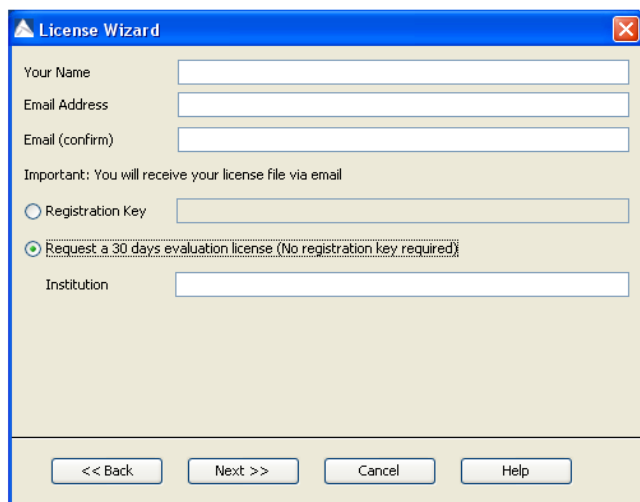
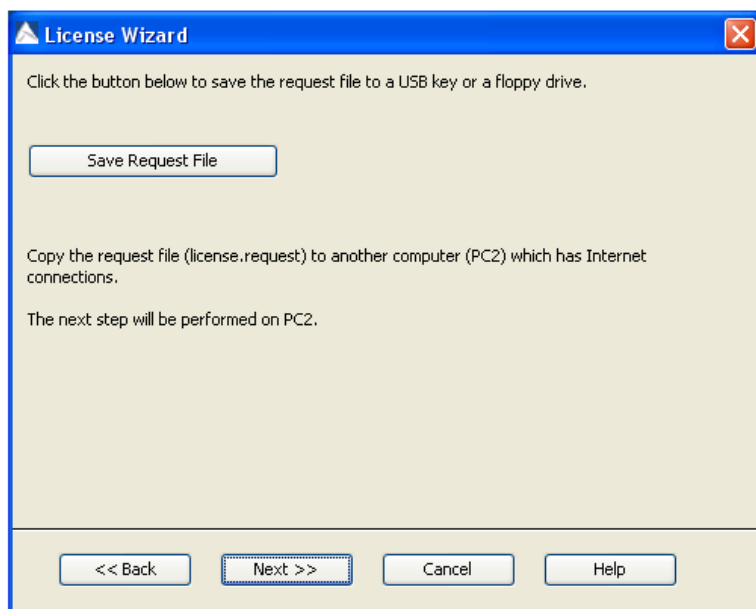
1. Select “Request license file (without Internet connection)” and click “Next” twice.

2. The following window will open:

If you have purchased PEAKS and have a registration key, select “Registration Key”. Enter your registration key as well as your name and email address and click “Next”.

OR

If you are trying a demo of PEAKS and do not have a registration key, select “Request a 30 days evaluation license (No registration key required)”. Enter your name, email address, as well as your institution. Click “Next”.

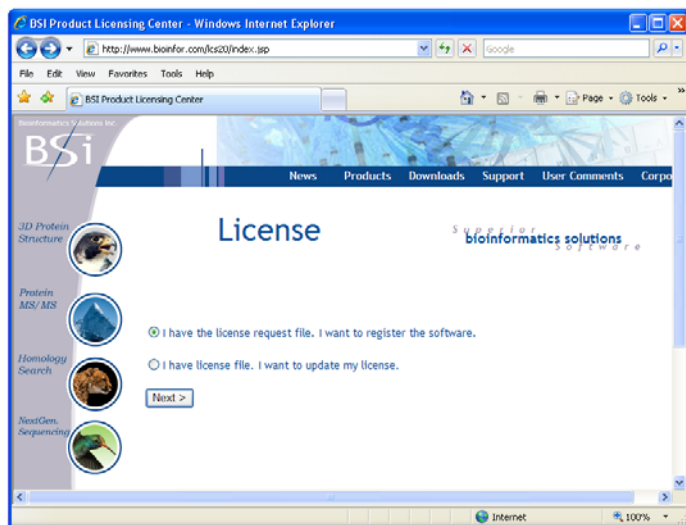
A Windows-style dialog box titled "License Wizard". It contains several input fields: "Your Name", "Email Address", "Email (confirm)", and "Institution". There are two radio button options: "Registration Key" and "Request a 30 days evaluation license (No registration key required)". The second option is selected. At the bottom, there are four buttons: "<< Back", "Next >>", "Cancel", and "Help".A Windows-style dialog box titled "License Wizard". It contains a button labeled "Save Request File". Below the button, there is text: "Click the button below to save the request file to a USB key or a floppy drive.", "Copy the request file (license.request) to another computer (PC2) which has Internet connections.", and "The next step will be performed on PC2.". At the bottom, there are four buttons: "<< Back", "Next >>", "Cancel", and "Help".

3. The following window will appear:

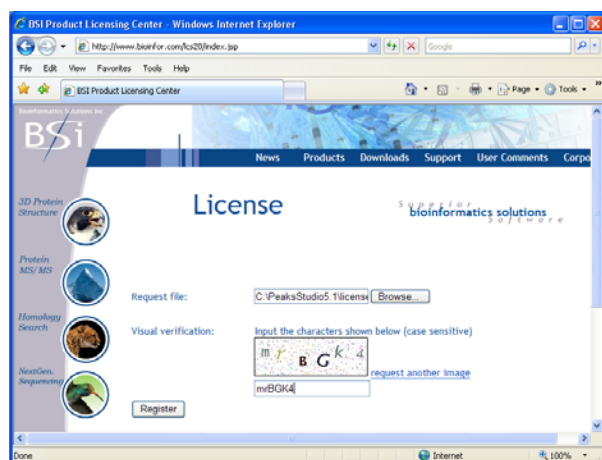
Select the “Save Request File” button to save license.request to your computer (PC1). Click “Next”.

4. Transfer the license.request file from PC1 onto a computer with an Internet connection (PC2) using a USB key or a removable storage device. On PC2, go to <http://www.bioinform.com/lcs20/>

5. Select “I have the license request file. I want to register the software” and click “Next”.

A screenshot of a web browser showing the "BSi Product Licensing Center" page. The page has a navigation bar with links: "News", "Products", "Downloads", "Support", "User Comments", and "Corporate". The main content area is titled "License" and contains two radio button options: "I have the license request file. I want to register the software." (which is selected) and "I have license file. I want to update my license.". Below these options is a "Next >" button. The left sidebar contains links to "3D Protein Structure", "Protein MS/MS", "Homology Search", and "NextGen Sequencing".

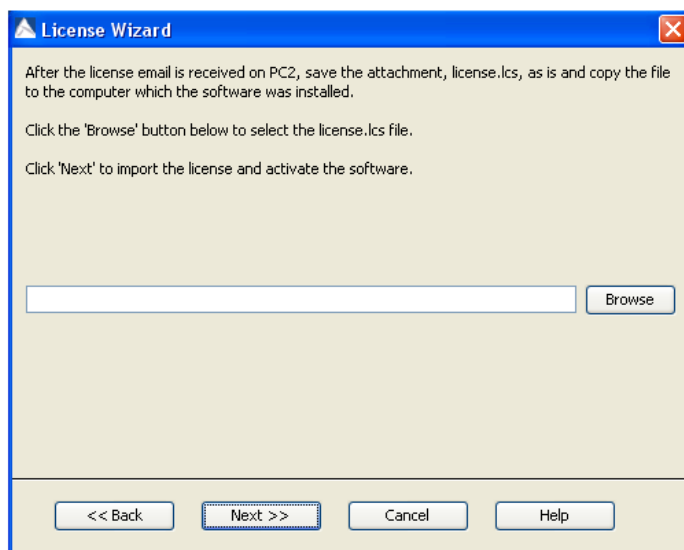
6. The following window will appear:



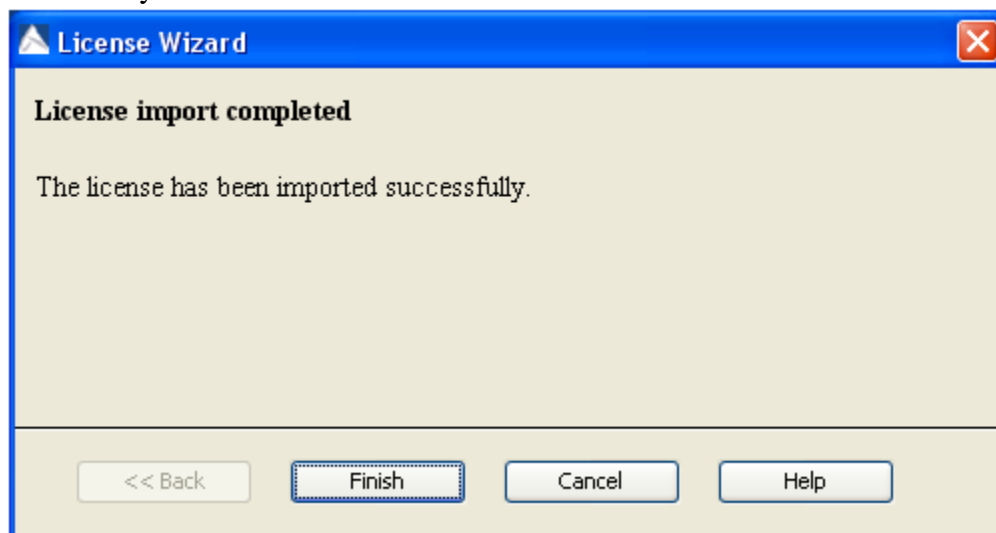
Click the “Browse” button to select the license.request file, type in the visual verification code and click “Next”.

7. After the license email is received on PC2, save the attachment, license.lcs, as is and copy the file to PC1. If you do not receive the license.lcs file in your inbox, please check your junk mail folder.

8. In the license wizard on PC1, click the 'Browse' button below to select the license.lcs file and click “Next”.



9. Click “Finish” if you receive a message that the license has been imported successfully.

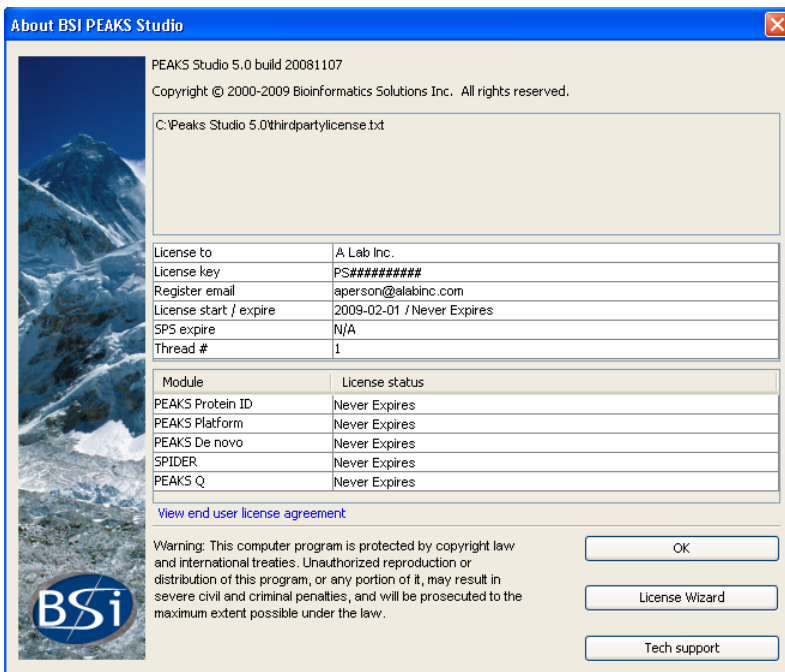


10. Restart PEAKS to begin using the software.

2.4.3 Re-registration Instructions

Re-registering PEAKS may be necessary if your license has expired or if you wish to update the license. You will need to obtain a new registration key from BSI. Once you have obtained this new key, select “About PEAKS Studio” from the Help menu. The “About BSI PEAKS Studio” dialogue box will appear:

Click the “License Wizard” button to continue. Then follow the instructions listed above for registering PEAKS.



2.5 Set up PEAKS Preferences

Before running your data, you must set up search engine preferences. For an explanation on how to do this, see section [15.1.3 Search Engine Preferences](#).

It is possible to run your data through PEAKS without changing the other preferences as they have default settings. For more information on changing these default settings see section [15.1 PEAKS Environment Preferences](#).

2.6 Set up PEAKS Configuration

Before running your data, you must configure your databases. For instructions on how to do this, see section [15.2.3 Database Configuration](#).

It is possible to run your data through PEAKS without configuring any other parameters as they have default settings. For more information on changing these default settings see section [15.2 PEAKS Configuration](#).


3. Quick Walkthrough

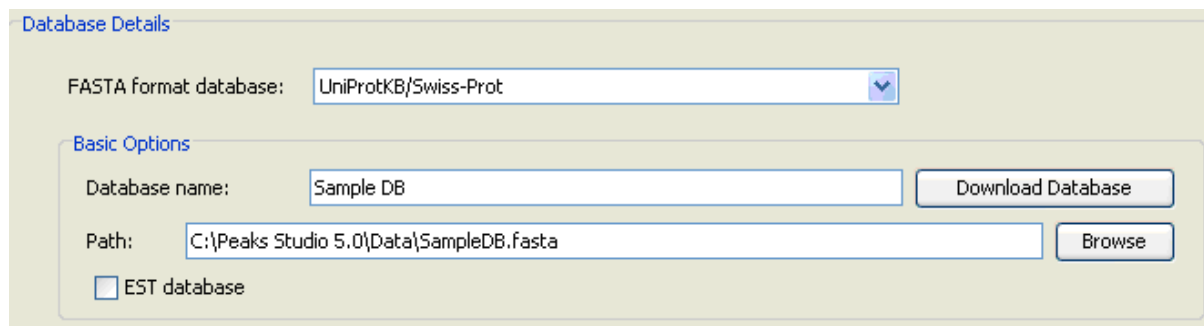
This section of the manual will walk you through most of the basic functionality of PEAKS 5.1. After completing this section you will see how easy it is to load, view a data file, perform data refinement, perform *de novo* sequencing, and database search protein identification.

Please note that version upgrades of PEAKS as well as upgrades to the databases may result in small changes to the results screenshots in this chapter.

3.1 Create a Sample Database


Before running the walkthrough data, you need to set up a database. So that this can be a quick process, we have provided you with a sample FASTA database called “SampleDB.fasta” in your PEAKS program folder (C:\PEAKSStudio5.1\Data).

Click on the configuration toolbar icon  or select “Configuration” from the “Tools” menu. Select “Database” from the left hand side of the window. Under “Database Details” enter the following information:

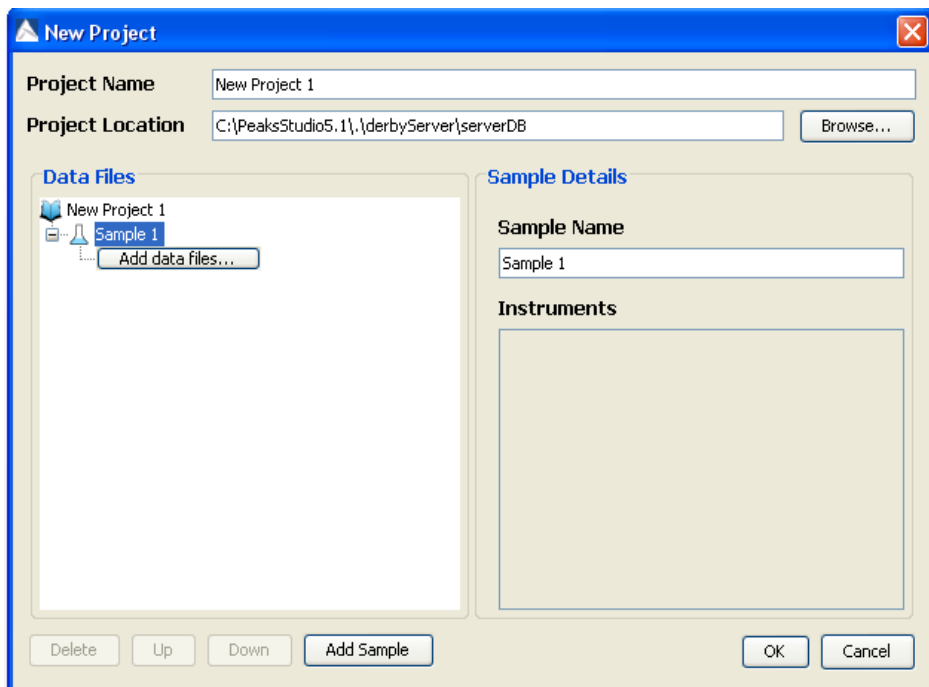


You do not need to change any of the other information listed. Click the “Add/Update” button and then click “OK”.

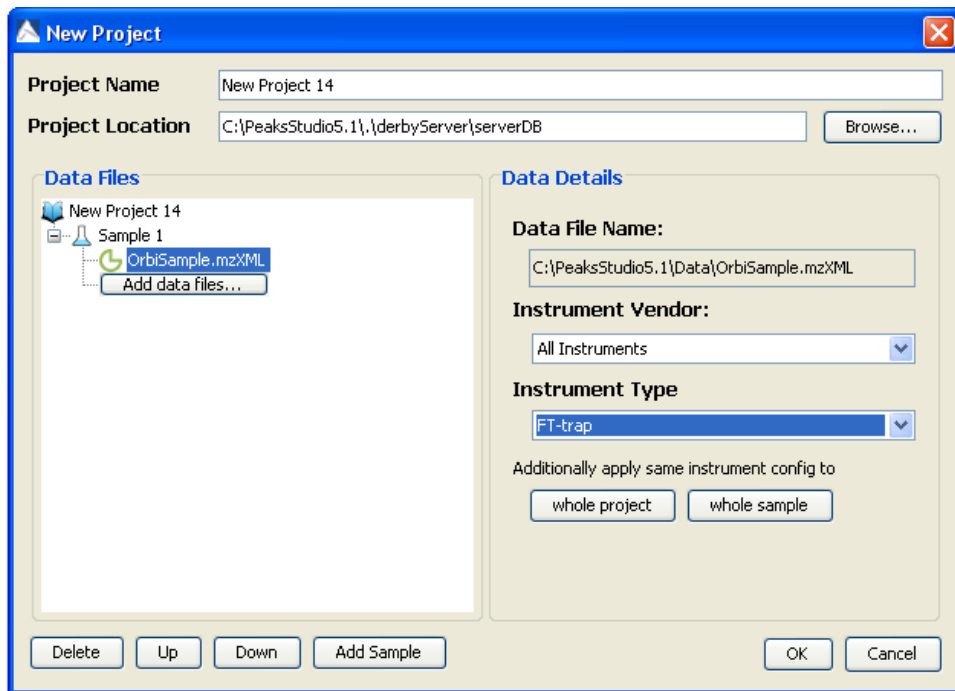
3.2 Create a Project

This will be a rather simple project as it will only contain one sample; however the same process will be used for projects with multiple samples and files. Click on the “Create new project” button  or select “New Project” from the “File” menu. The following window will appear:

Enter a name for your project. Do not change the default project location. By default the first sample will be named “Sample 1”. Click the “Add data files” button to add a data file for sample 1. Select the “OrbiSample.mzxml” file from your PEAKS program folder (for example



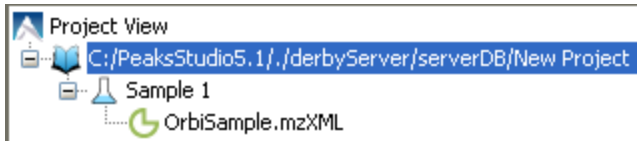
“C:\PEAKSStudio5.1\Data”) and click “Open”. Leave the instrument vendor as “All Instruments” and select “FT-trap” from the “Instrument Type” drop-down list.



Click “OK”.

You should now see the following in the “Project View” panel, which is located in the upper left hand corner of the screen:

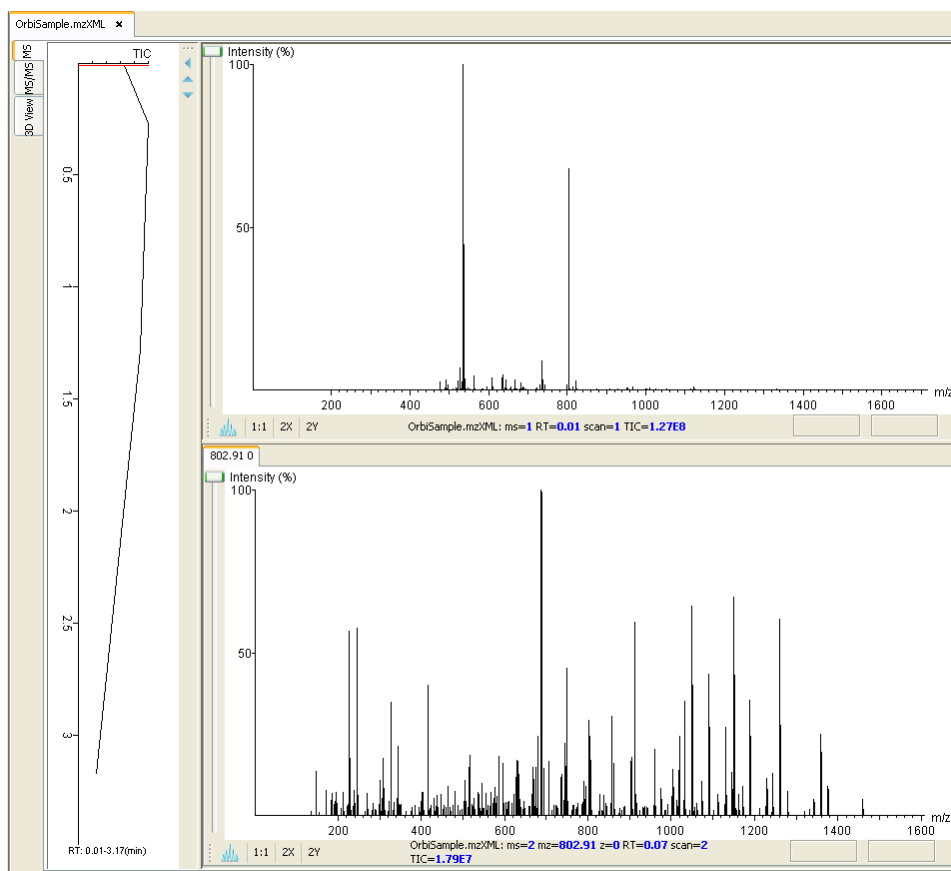
Use the ‘+’ and ‘-’ boxes to expand and collapse the project in order to access the data file in the “Project View”. Make sure that you select this data file when choosing data to be analyzed.



PEAKS reads and tracks information about the experiment and data for use in the analysis and for future reference. Once the data file has loaded, the fraction icon will turn solid. Click on the “Properties” tab in the bottom left hand corner:


Tasks	Running Info	Work Flow Info	Properties
Selection Details: OrbiSample.mzXML			
Total MSI Spectra	4		
Total MS/MS Spectra	6		
Ion Source	ESI(nano-spray)		
Fragmentation Mode	CID, CAD, IRMPD (y and b ions)		
MS Scan Mode	FT-ICR/Orbitrap		
MS/MS Scan Mode	Linear Ion Trap		
MS Scan Centroid	false		
MS/MS Scan Centroid	false		

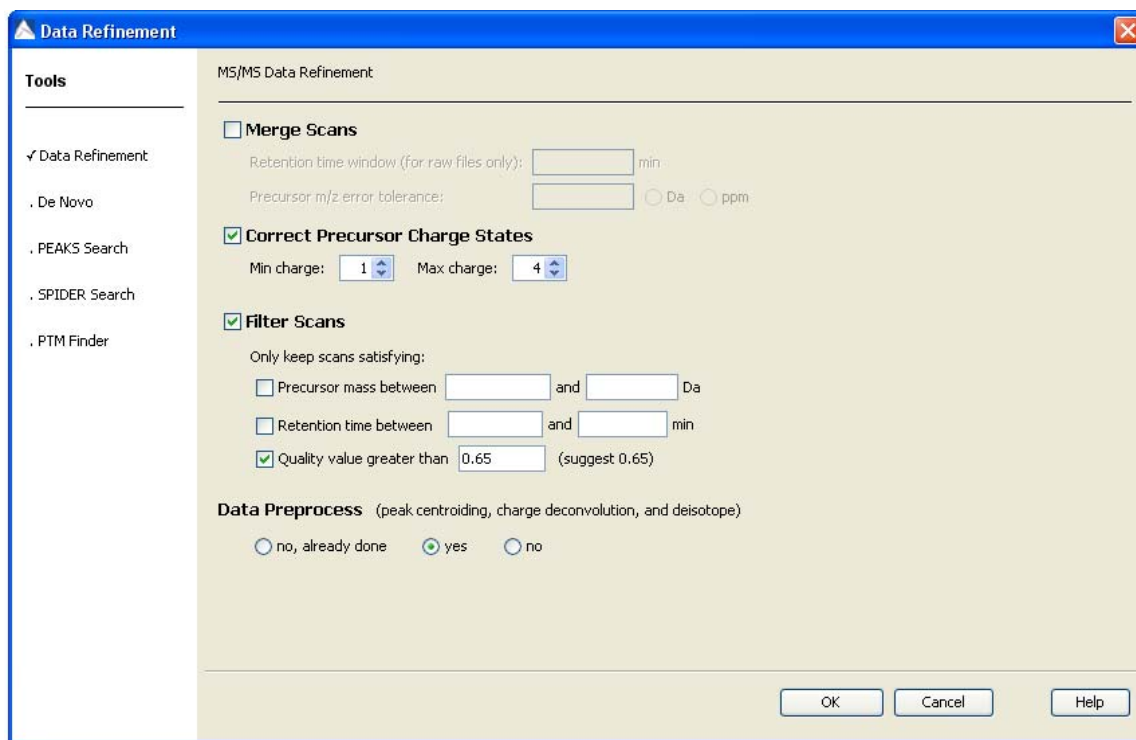
You should see the following in the “Main Processing Window”:



The information that is displayed by default pertains to the precursor scan. To the left of the window is the “Total Ion Chromatogram” (TIC). The graph in the upper right corner displays a survey scan with its corresponding tandem scans below. Click on the MS/MS tab to see the graphs that were generated from the tandem scan. For more information on the functions and tools found in these windows, see [section 4.8 Orienting Yourself](#).

3.3 Perform Data Refinement

- 1) Click the Data Refine toolbar icon 
Or
Select “Data Refine” from the “Tools” menu.
- 2) Enter the settings as shown:




For more details on setting up the parameters for data refinement refer to 5.2 Data Refinement Parameters.

Here we will use a quality filter to remove data with a quality value lower than 0.65. As all of the data in this data set is of good quality, we will not end up removing any data using this filter.

After running data refine, there will be new information listed in the “Properties” file.

Tasks		Running Info	Work Flow Info	Properties
Selection Details: OrbiSample.mzXML				
Total MSI Spectra	4			
Total MS/MS Spectra	6			
Ion Source	ESI(nano-spray)			
Fragmentation Mode	CID, CAD, IRMPD (y and b ions)			
MS Scan Mode	FT-ICR/Orbitrap			
MS/MS Scan Mode	Linear Ion Trap			
MS Scan Centroid	false			
MS/MS Scan Centroid	false			
Group MS/MS Scan (after refine)	6			
Charge Options	[1 - 4]			
Filter Quality	>0.65			
Process	true			

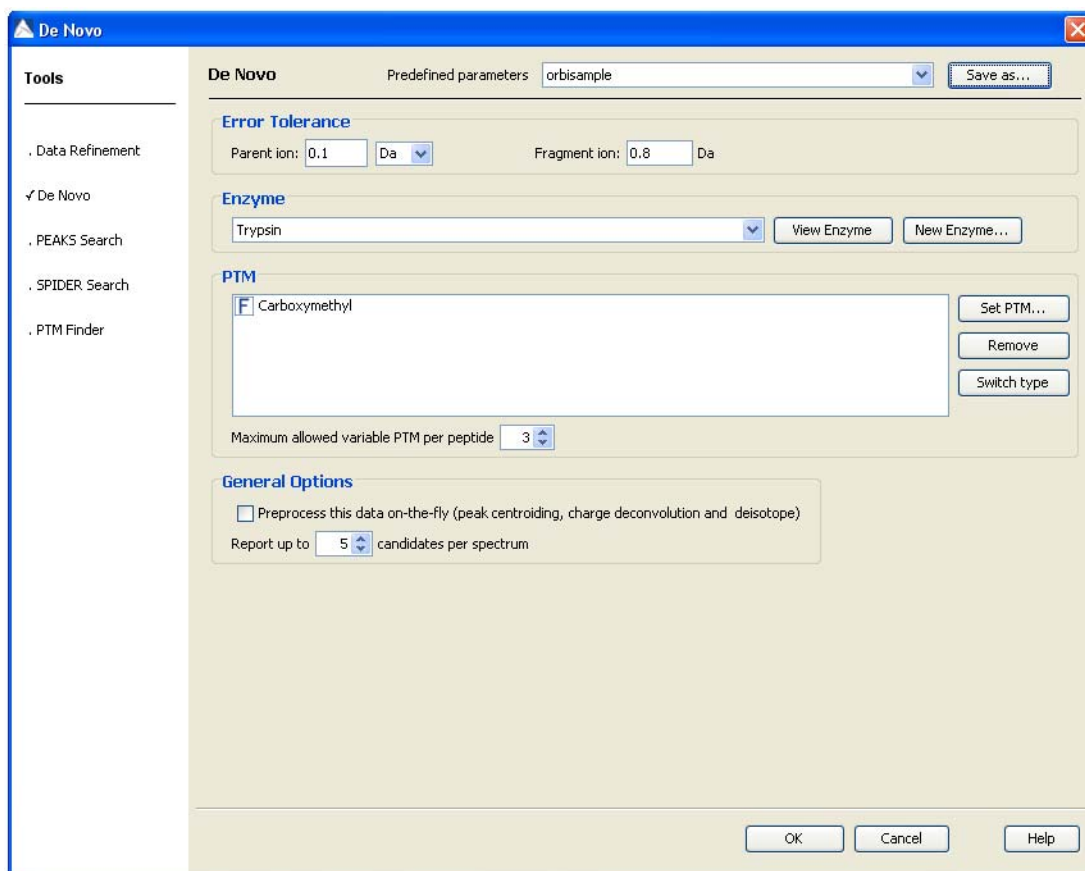
3.4 Run *de novo* Sequencing

1) Click the *de novo* sequencing toolbar icon 

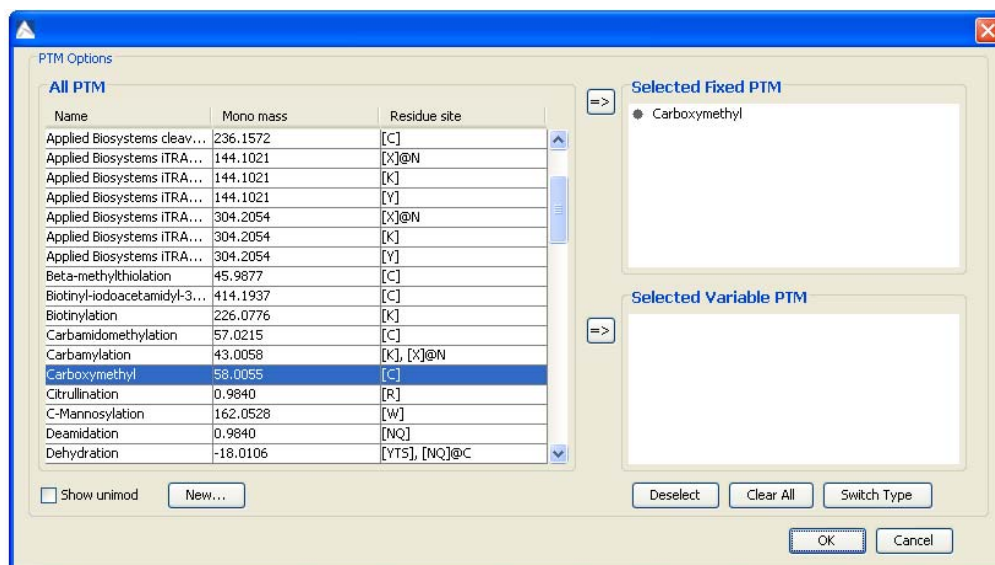
Or

Select “*de novo*” from the “Tools” menu.

2) Enter the settings as shown:



Enter the PTM as shown below by selecting “Set PTM”:



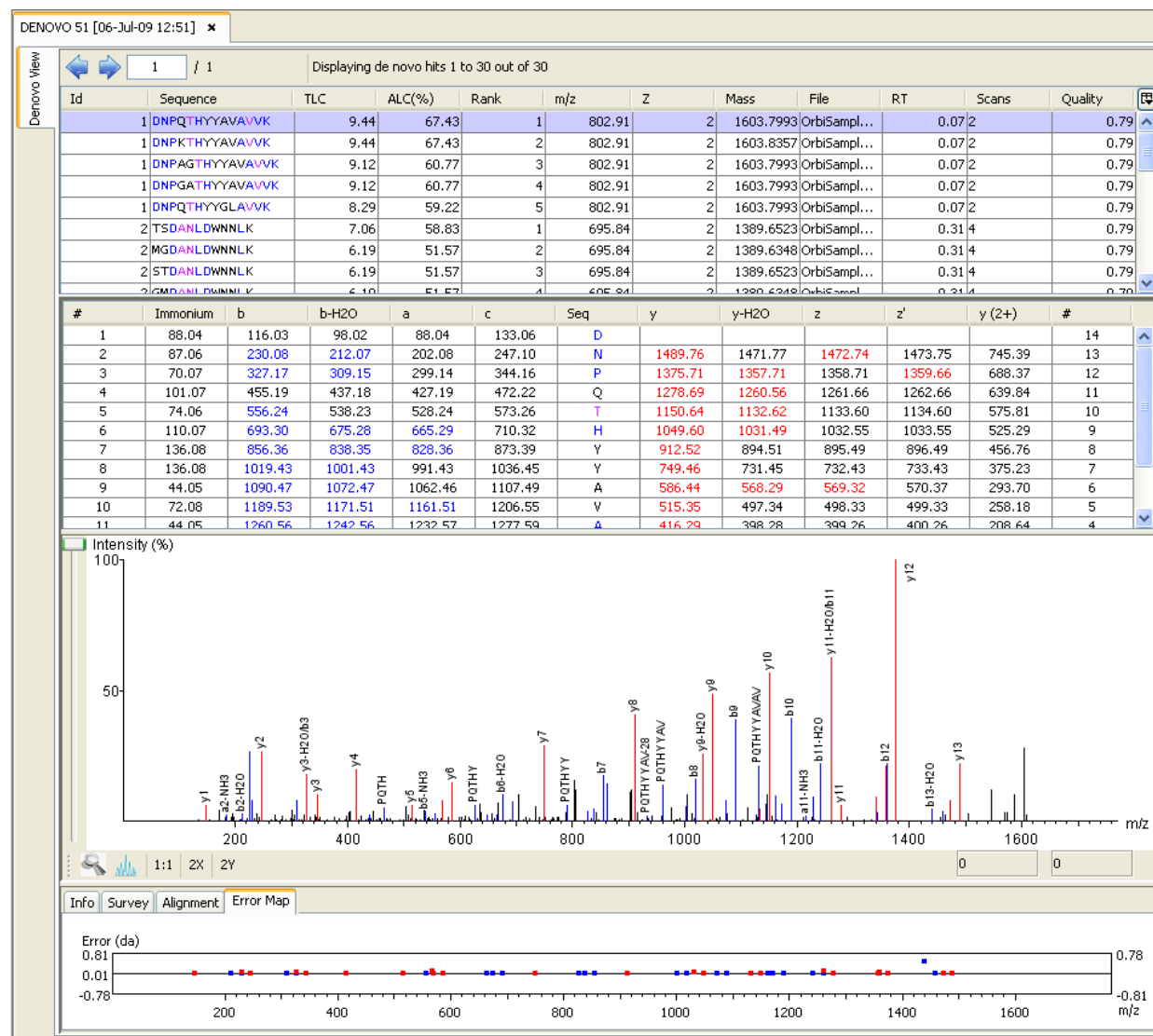
Note that we are not going to preprocess this data “on the fly” as we have already preprocessed the data during the data refinement stage. We will also choose to report only one peptide per spectrum for simplicity’s sake.

You can save the parameters that you used for future reference by clicking “Save As”. For more information on setting up *de novo* parameters see section [6.1 Setting up Auto *de novo* Sequencing Parameters](#). Click “OK” to commence analysis. The PEAKS auto *de novo* algorithm

derives sequence candidates for each of the six spectra in our example data file. Once the *de novo* sequencing is completed the results file will appear in the “Project View Panel”:



Double click on the *de novo* file and the results will appear in the “Main Processing Window”:



Take a look at spectra ID 6. Notice that the [1] refers to the modification which in this case is Carboxymethyl.

At the top of the screen you will see the peptide candidates in the “Peptide Candidates Frame”. The peptide candidates are sorted by “ID”. Right next to the proposed sequence, the auto *de*

novo “Total Local Confidence” (TLC) and “Average Local Confidence” (ALC%) confidence scores are shown. You will also see the *m/z* ratio, mass, retention time etc. listed for each peptide sequence. For information on the “Peptide Candidates Frame including *color* coding, see section [6.2.1 Peptide Candidates Frame](#).

Below the “Peptide Candidates Frame” is the “Ion Table Frame”. Each amino acid is *color* coded by confidence level (see section [6.2.2 Ion Table Frame](#)) with the masses for matched a, b and c ions listed in blue and for the matched x, y and z ions listed in red.

Below the “Ion Table Frame” is the “Spectrum View Frame”. This frame is useful for seeing the strength of the ms/ms peaks that PEAKS has set as ions. Here the alignment of the assigned b (blue) and y (red) ions with the entire spectrum corresponding to the selected peptide can be observed. For more information on the “Spectrum View Frame”, see section [6.2.3 Spectrum View Frame](#).

By clicking on the tabs at the bottom of the page, you can view the corresponding survey scan, the spectrum alignment and the error map. For more information, see section [6.2.4 Spectrum Alignment Frame](#), [6.2.5 Survey Scan](#) and [6.2.6 Error Map](#).

3.5 Run Protein Identification

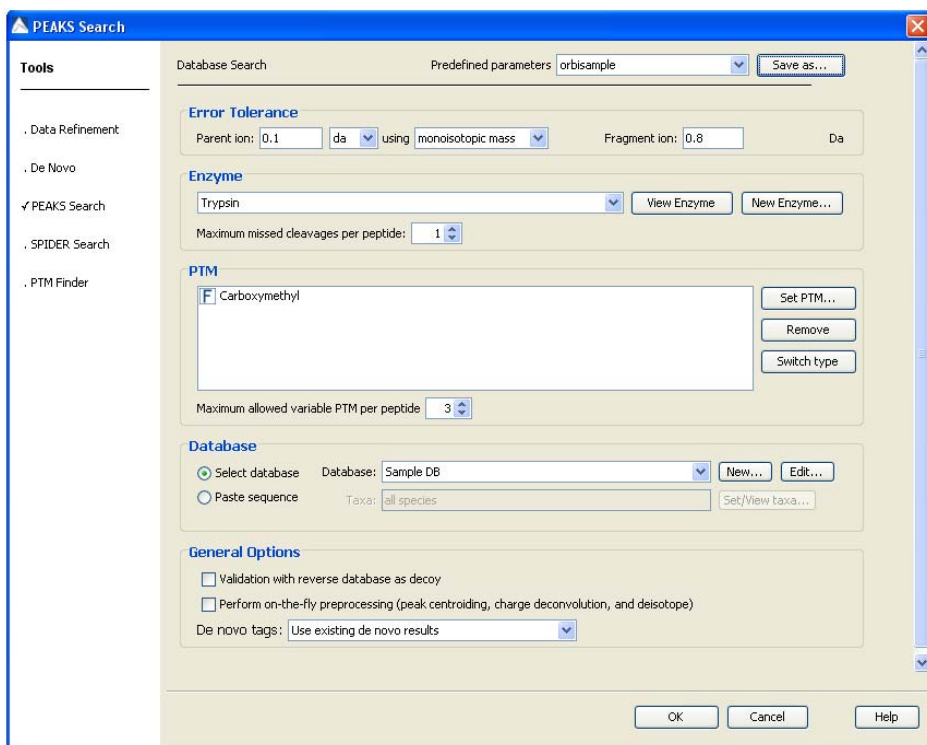
- 1) Click the PEAKS Search toolbar icon 

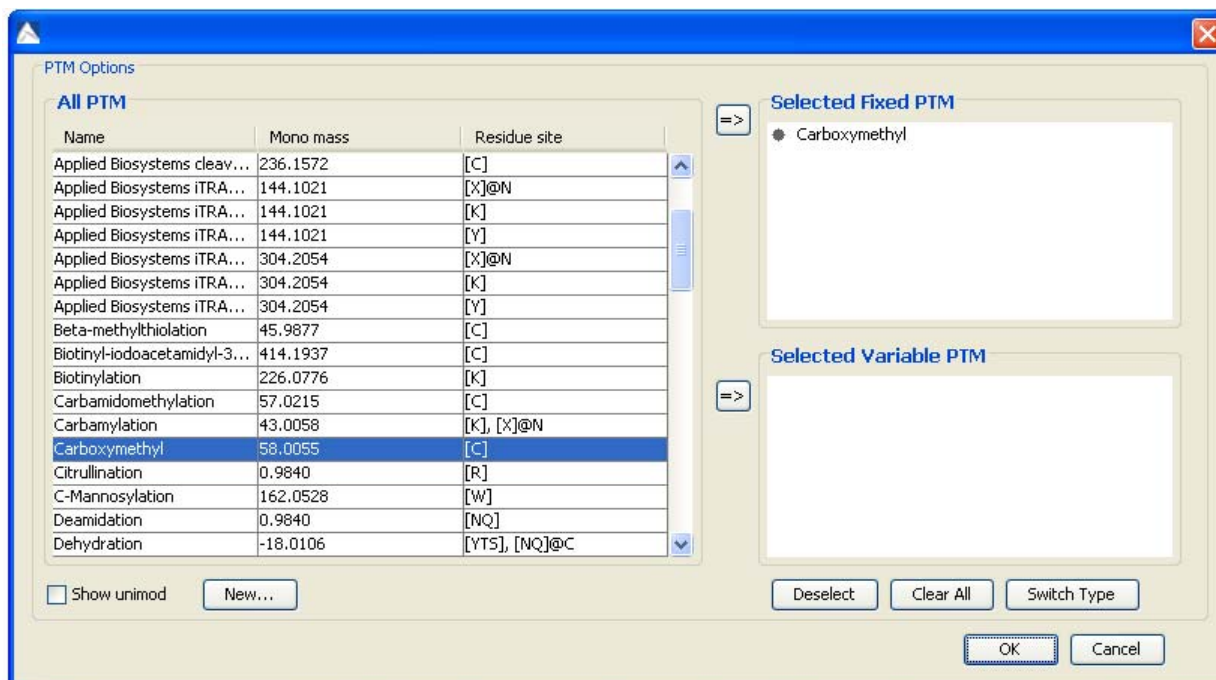
Or

Select “PEAKS Search” from the “Tools” menu.

- 2) Enter the settings as shown:

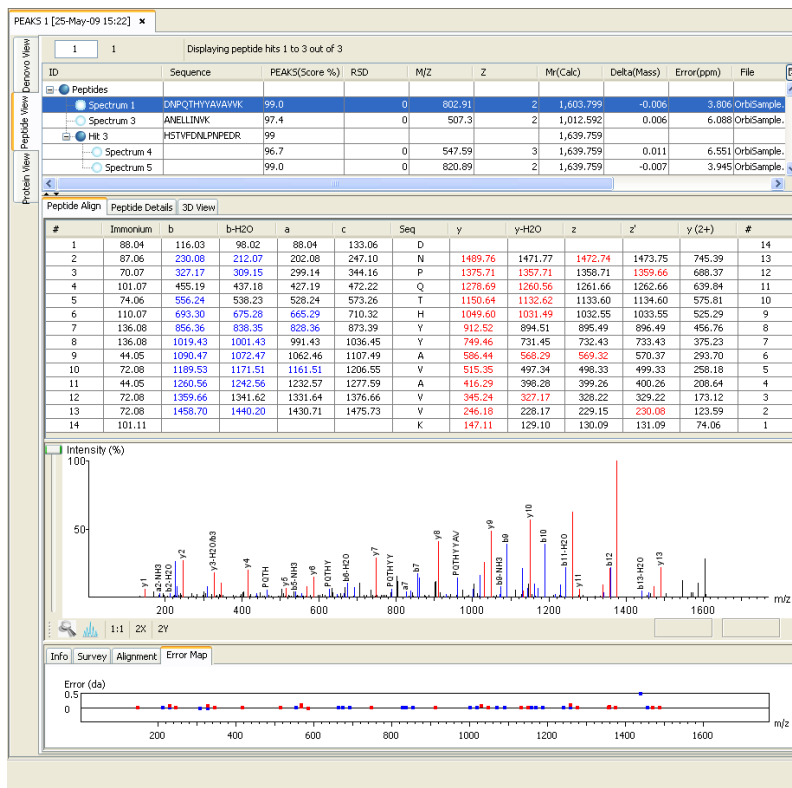
Enter the PTM as shown below by selecting “Set PTM”:





Parameters can be saved and used for future reference by clicking on the “Save Parameter” button. For more information on setting up protein identification parameters see section [7.1 Setting up Protein Identification Parameters](#).

Click “OK” to commence analysis.



Once the Protein ID file appears in the “Project View Panel”, double click on the file and the results will appear in the “Main Processing Window”:

The “Peptide View” will appear in the “Main processing” window by default. The “Peptide ID List” shows each spectrum for which PEAKS found a matching peptide. Since there may be more than one spectrum that matches a peptide, these spectra would be listed together under a Hit node. Use the ‘+’ and ‘-’ boxes to expand and collapse the node to see the spectra that are listed together. With this dataset, spectra 4 and 5 can both be found under one hit.

The “Peptide Alignment Panel” contains an “Ion Table”,

“Spectrum View Frame”, “Spectrum Alignment Frame”, “Survey Scan” and “Error Map”. Click on the “Peptide Details” tab to see a simple alignment between the original *de novo* sequence, the peptide found in the database and the reconstructed sequence. At the bottom of the “Peptide Details” panel you will see where the selected peptide matches the protein. For more information about these sections see sections [6.2.2 Ion Table Frame](#) - [6.2.6 Error Map](#) .

Click on the “Protein View” tab on the upper left hand side. The following window will appear:

Accession	ID	Mass	Display	PEAKS(Score %)	Coverage(%)	Query matched	Description	Marked
Q29443 TRFE_BOVIN	1	77753.200		83.8	3.69	3	Serotransferrin - Bos taurus (Bovine).	<input type="checkbox"/>
P00330 ADHI_YEAST	2	36691.957		59.3	2.31	1	Alcohol dehydrogenase 1 - Saccharomyces cere...	<input type="checkbox"/>

ID	Sequence	PEAKS(Score %)	RSD	M/Z	Z	Mr(Calc)	Delta(Mass)	Error(ppm)	File	RT	Scan	Quality	Start	End
Spectrum 1	DNPQTHYYAVVVK	99.0	0.0	802.91	2	1603.7993	-0.0061	3.8056	OrbiSample.mzXML	0.07	2	0.788	107	120
HR 3	HSTVFDNLNPEDR	99				1639.759							231	244
Spectrum 4		96.7	0.0	547.59	3	1639.759	0.0107	6.5511	OrbiSample.mzXML	1.34	7	0.777		
Spectrum 5		99.0	0.0	820.89	2	1639.759	-0.0065	3.9455	OrbiSample.mzXML	1.42	8	0.787		

PEAKS 5.1 presents a list of proteins, ranked in descending order from highest score on downward. Clicking on any protein will display the peptides matched to that protein in the bottom pane. In this case the highest scoring protein match is Serotransferrin precursor from bovine. This protein has two matching peptides, which you can see in the “Peptide ID” list when if you have selected the “Peptides” tab. To see exactly where the peptide corresponds to the protein sequence, select the “Coverage” tab. The entire sequence of the protein is shown and the matched parts are highlighted in blue. In this case the total matched part accounts for 3.69% of the protein. This information can be found in the “Coverage” column.

Peptides	Coverage	Tool Box
DB peptides shown in blue		
1	MRPAVRALLA CAVLGLCLAD PERTVRWCTI STHEANKCAS FRENVLRILE	
51	SGPFVSCVKK TSHMDCIKAI SNNEADAVTL DGGLVYEAGL KPNNLKPVVA	
101	EFHGTDNPQ THYYAVVVK KDTDFKLNEL RGKKSCHTGL GRSAGWNIPM	

3.6 Run PTM Finder

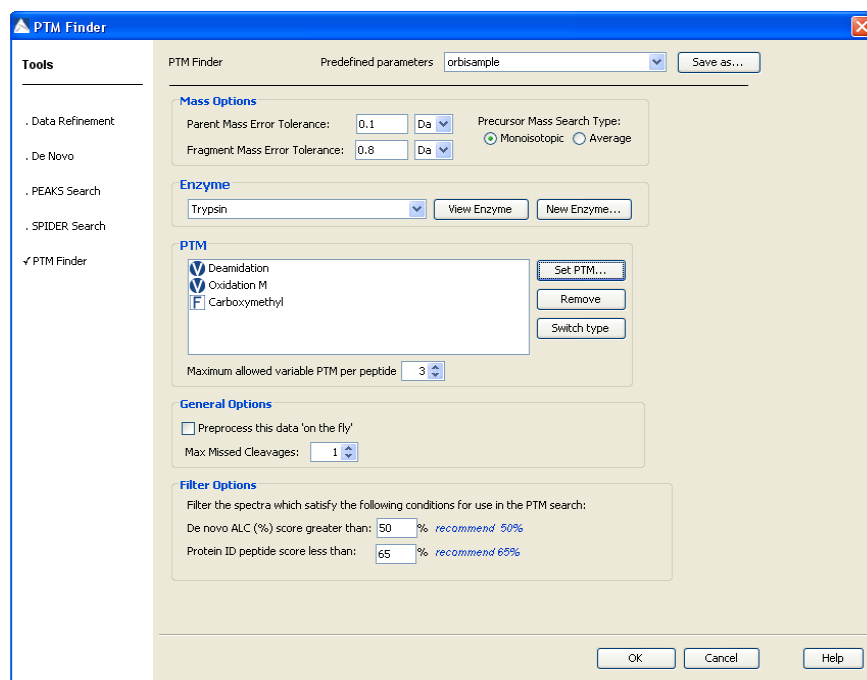
Using the PTM finder, you can identify any additional PTMs and increase the coverage of the proteins that we have found. It is important to note that the PTM Finder can only be applied to a Protein ID results file. As it is very time consuming to run Protein ID with many PTMs, this allows searching for more PTMs in less time. Make sure that you click on a Protein ID result before performing a PTM Finder search.

- 1) Click the PTM Finder toolbar icon

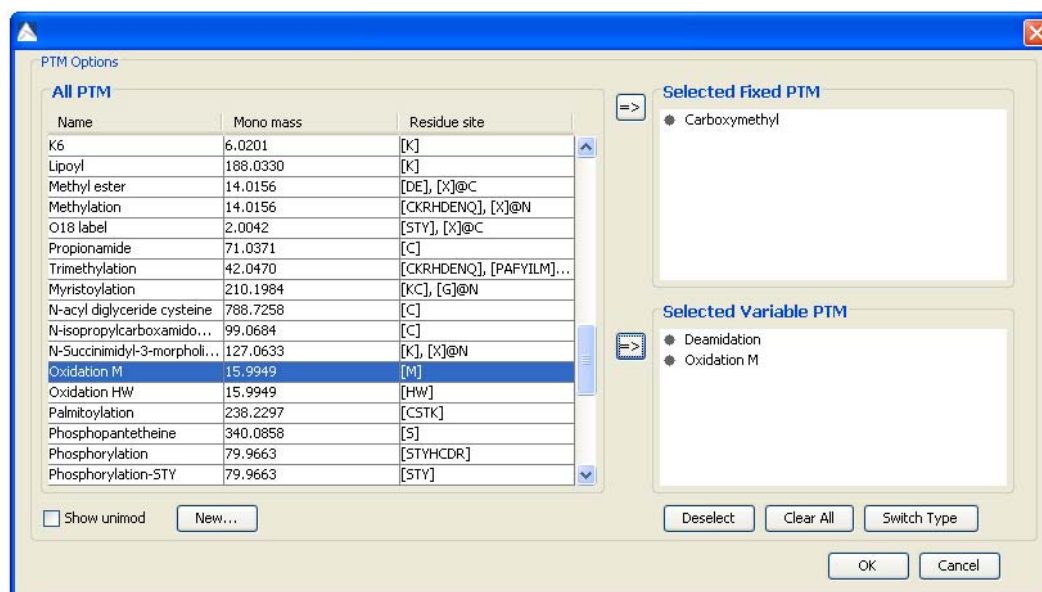
Or

Select “PTM Finder” from the “Tools” menu.

- 2) Enter the settings as shown:



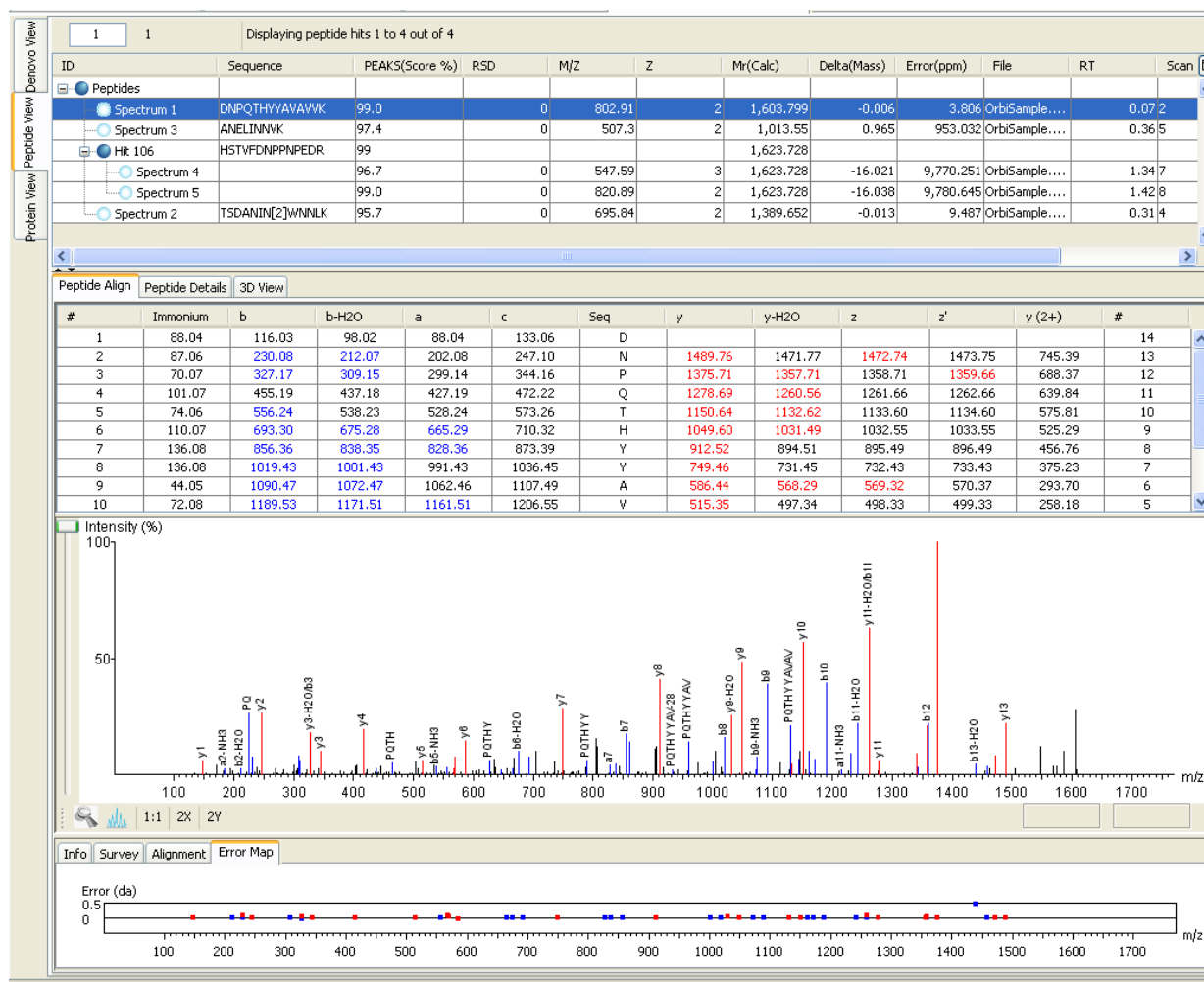
Enter the PTM as shown below by selecting “Set PTM”:



Saving the parameters for future reference is achieved by clicking on the “Save Parameter” button. For more information on setting up PTM Finder parameters see [section 9.1 Setting up PTM](#)

[Finder Parameters](#). Click “OK” to commence analysis.

After the PTM Finder search is complete, double click on the result node. The “Peptide View” window will appear:



The inChorus search window will open. Check the “PEAKS Protein ID” box and select the name. Enter the following settings:

inChorus Search

Tools

- ☒ PEAKS Protein ID
- ☐ X!Tandem
- ☐ OMSSA
- ☐ Mascot
- ☐ Sequest
- ☐ Import Result

PEAKS Protein ID

Database Search: Predefined parameters: orbisample [Save as...]

Error Tolerance

Parent ion: 0.1 da using monoisotopic mass Fragment ion: 0.8 Da

Enzyme

Trypsin [View Enzyme] [New Enzyme...]

Maximum missed cleavages per peptide: 1

PTM

Deamidation [Set PTM...]
 Oxidation M [Remove]
 Carboxymethyl [Switch type]

Maximum allowed variable PTM per peptide: 3

Database

Select database Database: Sample DB [New...] [Edit...]
 Paste sequence Taxa: all species [Set/View taxa...]

General Options

☐ Validation with reverse database as decoy
☐ Perform on-the-fly preprocessing (peak centroiding, charge deconvolution, and deisotope)
 De novo tags: Use existing de novo results

[OK] [Cancel] [Help]

Enter the PTM as shown below by selecting “Set PTM”:

PTM Options

All PTM

Name	Mono mass	Residue site
K6	6.0201	[K]
Lipoyl	188.0330	[K]
Methyl ester	14.0156	[DE], [X]@C
Methylation	14.0156	[CKRHDENQ], [X]@N
O18 label	2.0042	[STY], [X]@C
Propionamide	71.0371	[C]
Trimethylation	42.0470	[CKRHDENQ], [PAFYILM]...
Myristoylation	210.1984	[KC], [G]@N
N-acyl diglyceride cysteine	788.7258	[C]
N-isopropylcarboxamido...	99.0684	[C]
N-Succinimidyl-3-morpholi...	127.0633	[K], [X]@N
Oxidation M	15.9949	[M]
Oxidation HW	15.9949	[HW]
Palmitoylation	238.2297	[CSTK]
Phosphopantetheine	340.0858	[S]
Phosphorylation	79.9663	[STYHCDR]
Phosphorylation-STY	79.9663	[STY]

☐ Show unimod [New...]

Selected Fixed PTM

- Carboxymethyl

Selected Variable PTM

- Deamidation
- Oxidation M

[Deselect] [Clear All] [Switch Type]

[OK] [Cancel]

Next, check the “X!Tandem” box.

Enter the following settings:

InChorus Search

Tools

- ☐ PEAKS Protein ID
- ☒ **X!Tandem**
- ☐ OMSSA
- ☐ Mascot
- ☐ Sequest
- ☐ Import Result

X!Tandem

- Database**

Sequence library: Taxonomy (ctrl key for multiple selection):

Taxonomy supported

 - all species
 - Archaea
 - Aeropyrum pernix
 - Methanococcus jannaschii
 - Other Archaea
 - Bacteria
 - Actinobacteria

1. ☐ reversed sequences

2. ☐ all ¹⁵N Find models with log(e) <
- measurement errors**

1. ☒ Fragment mass error:
- residue modifications**

1. Complete modifications:

☒ specify your own

2. Potential modifications:

☒ specify your own
- refinement specification**

1. ☒ Potential modifications (unimod):

round 1	round 2
none	none
Oxidation (M)	Oxidation (M)
Oxidation (W)	Dioxidation (M)
Deamidation (N)	Oxidation (W)
Deamidation (Q)	Dioxidation (W)
ICAT-D:2H(8) (C)	Deamidation (N)
mods: <input type="text"/>	mods: <input type="text"/>
motifs: <input type="text"/>	motifs: <input type="text"/>

2. ☒ Point mutations: ☐ yes ☒ no

3. ☒ Semi-style cleavage: ☐ yes ☒ no
- protein cleavage specification**

1. ☒ Cleavage site:

2. ☒ Semi-style cleavage: ☐ yes ☒ no
- spectrum conditioning**

1. ☒ Remove redundant: ☐ yes ☒ no, angle: (0-90)

2. ☒ Spectrum synthesis: ☒ yes ☐ no
- predefined methods**

1. ☒ Method: Select device & parent δm .

 - FTICR (10 ppm)
 - Quad-TOF (100 ppm)
 - Quad-TOF (0.5 Da)
 - Ion Trap (4 Da)

Finally, check the “OMSSA” box and select the name.

Enter the following settings:

The screenshot shows the InChorus Search window with the OMSSA search engine selected. The interface includes a left-hand menu with search tools (PEAKS Protein ID, X!Tandem, OMSSA, Mascot, Sequest, Import Result) and a main search area. The OMSSA search area contains various configuration options:

- Enzyme:** Trypsin
- Maximum missed cleavages:** 3
- Sequence library:** Sample DB (Taxonomy supported)
- Species to search:** all species (includes Archaea, Bacteria, Actinobacteria)
- Hitlist max length:** 10
- E-value cutoff:** 1
- Fixed mods:** carboxyamidomethylation of K, carboxykynurenin of W, carboxymethyl C, citrullination of R, deamidation of N and Q, dehydro of S and T, di-O18 on peptide n-term
- Variable mods:** citrullination of R, deamidation of N and Q, dehydro of S and T, di-O18 on peptide n-term, di-iodination of Y, di-methylation of K, oxidation of M
- Maximum variable mod combinations searched per peptide:** 64
- Precursor mass tolerance (Da):** 0.1
- Product mass tolerance (Da):** 0.8
- Precursor mass search type:** monoisotopic
- Product mass search type:** monoisotopic
- Lower bound of precursor charge:** 1
- Upper bound of precursor charge:** 3
- Minimum charge to start using multiply charged products:** 3
- Fraction of product peaks below precursor to determine +1 precursor:** 0.95
- Peak intensity cutoff:** 0 (fraction of most intense)
- Number of top intensity peaks in first pass:** 6
- Ions to search 1:** b
- Ions to search 2:** y

At the bottom of the window are buttons for OK, Cancel, and Help.

Click the “Ok button. When the inChorus search is complete you should see the following new additions in the “Project View” panel:

- PEAKS 1 [21-Jan-09 11:38]
- X! XTANDEM 2 [21-Jan-09 11:39]
- M OMSSA 3 [21-Jan-09 11:40]
- INCHORUS 4 [21-Jan-09 11:40]**

Presented here are individual reports for PEAKS, X!Tandem and OMSSA as well as an inChorus report that compares the individual reports. To see each of these reports, double click on the report that you would like to see in the “Project View” panel.

The “Peptide View” results for the PEAKS Protein ID search can be seen below:

PEAKS 4 [22-Jun-09 13:20] x

1 / 1 Displaying peptide hits 1 to 4 out of 4

ID	Sequence	PEAKS(Score %)	M/Z	Z	Mr(Calc)	Delta(Mass)	Error(ppm)	File	RT	Scan
Peptides										
Spectrum 1	DNPQTHYYAVAVVK	99.0	802.91	2	1,603.799	-0.006	3.806	OrbiSample....	0.07	2
Spectrum 2	TSDANIN[3]WNNLK	99.0	695.84	2	1,389.652	-0.013	9.487	OrbiSample....	0.31	4
Spectrum 3	ANELLINVK	97.5	507.3	2	1,012.592	0.006	6.088	OrbiSample....	0.36	5
Hit 13	HSTVFDNLPNPEDR	99			1,639.759					
Spectrum 4		96.7	547.59	3	1,639.759	0.011	6.551	OrbiSample....	1.34	7
Spectrum 5		99.0	820.89	2	1,639.759	-0.007	3.945	OrbiSample....	1.42	8

The “Peptide View” results for the X!Tandem search can be seen below:

XTANDEM 11 [22-Jun-09 14:16] x

1 / 1 Displaying peptide hits 1 to 3 out of 3

ID	Sequence	XTANDE...	M/Z	Z	Mr(Calc)	Delta(Mass)	Error(ppm)	File	RT	Scan
Peptides										
Spectrum 1	DNPQTHYYAVAVVK	1.6E-6	802.91	2	1,603.799	-0.006	3.806	OrbiSample....	0.07	2
Spectrum 2	TSDANIN[4]WNNLK	3.5E-2	695.84	2	1,389.652	-0.013	9.487	OrbiSample....	0.31	4
Spectrum 5	HSTVFDNLPNPEDR	7.7E-6	820.89	2	1,639.759	-0.007	3.945	OrbiSample....	1.42	8

“Peptide View” results for the OMSSA search can be seen below:

OMSSA 6 [22-Jun-09 13:21] x

1 / 1 Displaying peptide hits 1 to 6 out of 6

ID	Sequence	OMSSA(E-value)	M/Z	Z	Mr(Calc)	Delta(Mass)	Error(ppm)	File	RT	Scan
Peptides										
Spectrum 1	DNPQTHYYA...	4.14E-12	802.91	2	1,603.799	-0.006	3.806	OrbiSample....	0.07	2
Spectrum 2	TSDANIN[5]...	3E-3	695.84	2	1,389.648	-0.017	12.298	OrbiSample....	0.31	4
Spectrum 2	TSDAN[5]IN...	2.38E-1	695.84	2	1,389.648	-0.017	12.298	OrbiSample....	0.31	4
Spectrum 2	TSDANINW...	5.76E-1	695.84	2	1,389.648	-0.017	12.298	OrbiSample....	0.31	4
Spectrum 3	ANELLINVK	4.71E-2	507.3	2	1,012.592	0.006	6.088	OrbiSample....	0.36	5
Hit 23	HSTVFDNLP...	2.54E-7			1,639.759					
Spectrum 4		1.73E-1	547.59	3	1,639.759	0.011	6.551	OrbiSample....	1.34	7
Spectrum 5		2.54E-7	820.89	2	1,639.759	-0.007	3.945	OrbiSample....	1.42	8

The inChorus report contains most of the information that is seen in a PEAKS Protein ID results file (7.2 Protein Identification Results). Click on the “Peptide View” tab:

INCHORUS 6 [26-Jun-09 14:48] x

1 / 1 Displaying peptide hits 1 to 4 out of 4

ID	Sequence	M/Z	Z	RT	InChorus(Score %)	PEAKS(Score %)	XTANDEM(E-value)	OMSSA(E-value)
Peptides								
Spectrum 1	DNPQTHYYAVAVVK	802.91	2	0.07	99.98	99.0	1.5E-6	1.37E-12
Spectrum 2	TSDANIN[2]WNNLK	695.84	2	0.31	99.49	99.0	2.4E-2	-
Spectrum 3	ANELLINVK	507.3	2	0.36	99.86	97.5	-	2.57E-2
Hit 21	HSTVFDNLPNPEDR				99.98			
Spectrum 4		547.59	3	1.34	99.82	96.7	-	5.37E-2
Spectrum 5		820.89	2	1.42	99.98	99.0	6.9E-6	7.75E-8

The “Peptide View” of an inChorus report contains the scores received by each search engine involved in the inChorus search. A “-” indicates that the search engine did not find that a protein sequence for that particular spectrum. Notice that while PEAKS Protein ID found spectra 1-5, X!Tandem found 1, 5 and OMSSA found 1, 3, 4 and 5.


Click on the “Protein View” tab:

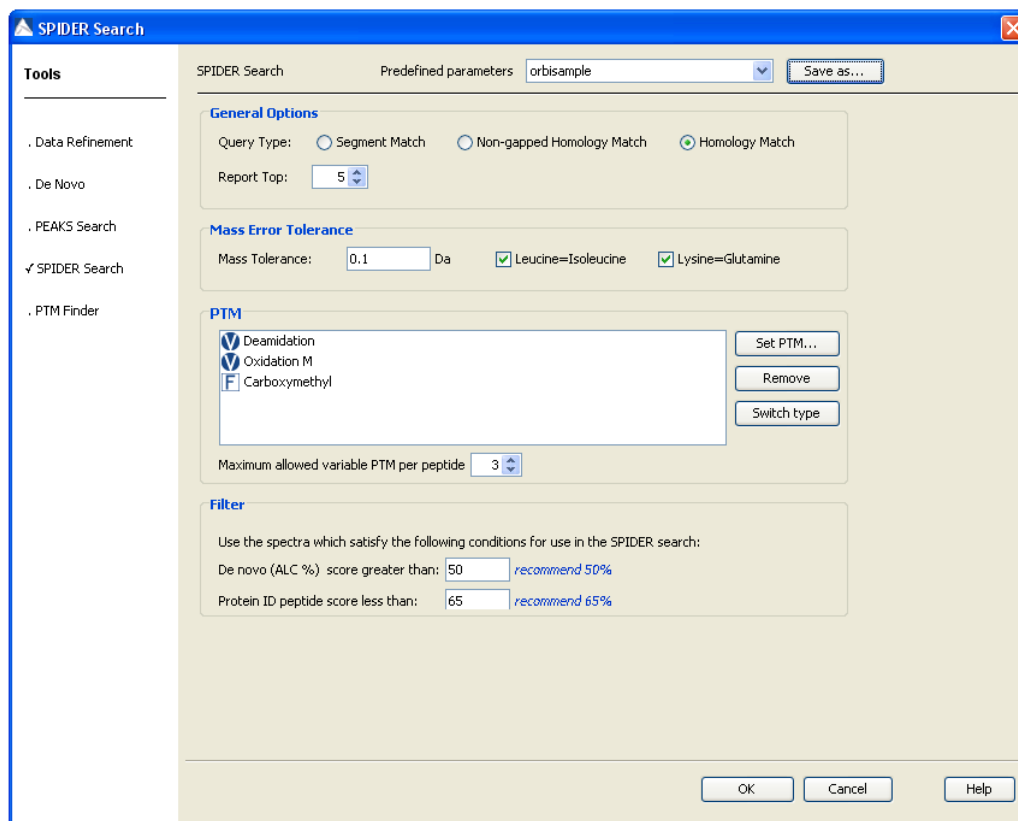
Accession	ID	Mass	Display	InChorus(Score %)	Coverage(%)	Query matched	Marked	Description	PEAKS	XTANDEM	OMSSA
Q29443 TRFE_BOVIN	1	77753.2		98.43	5.26	6		Serotransferin - ...	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
P00330 ADH1_YEAST	2	36691.957		61.53	2.31	1		Alcohol dehydrog...	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

The “Protein View” of an inChorus report displays the proteins that were found and indicates by checkmarks whether the different search engines found that protein or not. In this case, the X!Tandem and OMSSA searches did not generate any extra results that PEAKS did not find but helped to confirm that the first protein is good match.

3.8 Perform a SPIDER Search

In this example, spectrum 6 has not been identified with database searching tools. In order to gain more information from our data, we will run a SPIDER search next. For more information about the SPIDER search see section [8. SPIDER Search](#).

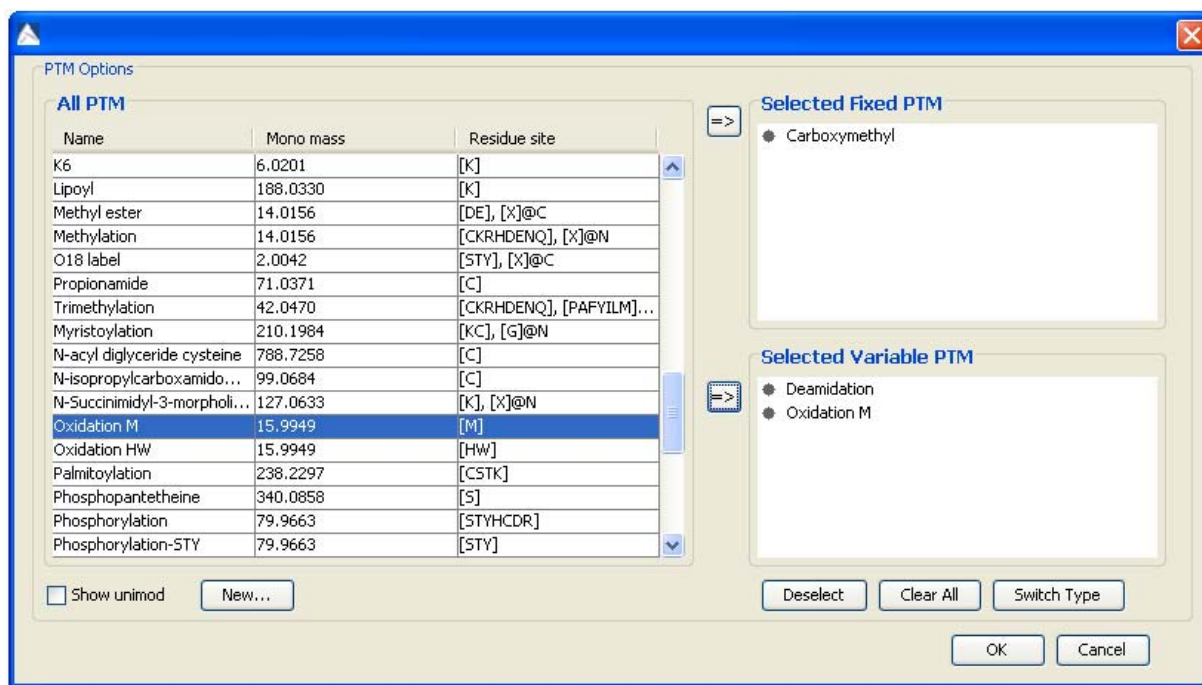
- 1) Click on the Protein ID Result
- 2) Next, click the SPIDER Search toolbar icon 
Or
Select “SPIDER Search” from the “Tools” menu.
- 3) Enter the settings as shown:



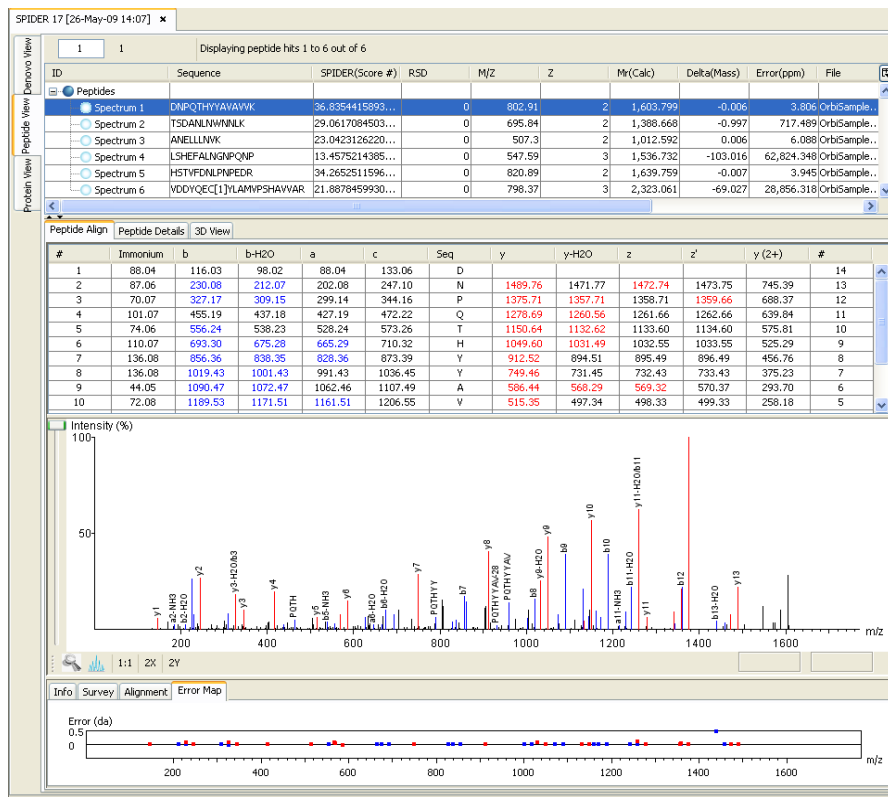
The SPIDER Search dialog box is shown with the following settings:

- Tools:** SPIDER Search is selected.
- Predefined parameters:** orbisample
- General Options:**
 - Query Type: ☒ Homology Match
 - Report Top: 5
- Mass Error Tolerance:**
 - Mass Tolerance: 0.1 Da
 - ☒ Leucine=Isoleucine
 - ☒ Lysine=Glutamine
- PTM:**
 - ☒ Deamidation
 - ☒ Oxidation M
 - ☐ Carboxymethyl
 - Maximum allowed variable PTM per peptide: 3
- Filter:**
 - De novo (ALC %) score greater than: 50 (recommend 50%)
 - Protein ID peptide score less than: 65 (recommend 65%)

Enter the PTM as shown below by selecting “Set PTM”:



Parameters can be saved for future reference by clicking on the “Save Parameter” button. For more information on setting up SPIDER Search parameters see section [8.1 Setting up SPIDER Parameters](#). Click “OK” to commence analysis.



After the SPIDER search has completed, double click on the SPIDER results node. The “Peptide View” window will appear. The format is identical to what was seen in the results of a Protein ID search:

Note that all spectra can now be identified by the SPIDER search. Spectrum 6 is well identified.

Clicking on the “Peptide Details” tab will display the protein with its matched peptides in red. SPIDER will also display a reconstructed sequence. See section [8.2 SPIDER Results View](#) for more information.

Peptide Align

Peptide Details

3D View

Select peptides for display:

Q29443|TRFE_BOVIN

Denovo

W (Carboxymethyl) W

<WFL>

Y L A M V P S H A { R L } R

Recon

AD (Carboxymethyl) W

<RE(Carboxymethyl)C>

Y L A M V P S H A { V V A } R

Homolog

VD

D

<QE(Carboxymethyl)C>

Y L A M V P S H A { V V A } R

SPIDER matches shown in red

1

MRPAVRALLA CAVLGLCLAD PERTVRWCTI STHEANKCAS FRENVLRLIE

51

SGPFVSCVKK TSHMDCIKAI SNNEADAVTL DGGLVYEAGL KPNNLKPVVA

101

EFHGTDKNPQ THYYAVAVVK KDTDFKLNEL RGKKSCHTGL GRSAGWNIPM

151

AKLYKELPDP QESIQRAAAN FFSASCVPCA DQSSFPKLCQ LCAGKGTDKC

201

ACSNHEPYFG YSGAFKCLME GAGDVAFAVKH STVFDNLPNP EDRKNYELLC

251

GDNTRKSVDD YQECYLAMVP SHAVVARTVG GKEDVIWELL NHAQEHFGKD

Peptides

Coverage

Tool Box

SPIDER matches shown in red

Overlapping matches in magenta

1

MRPAVRALLA CAVLGLCLAD PERTVRWCTI STHEANKCAS FRENVLRLIE

51

SGPFVSCVKK TSHMDCIKAI SNNEADAVTL DGGLVYEAGL KPNNLKPVVA

101

EFHGTDKNPQ THYYAVAVVK KDTDFKLNEL RGKKSCHTGL GRSAGWNIPM

151

AKLYKELPDP QESIQRAAAN FFSASCVPCA DQSSFPKLCQ LCAGKGTDKC

201

ACSNHEPYFG YSGAFKCLME GAGDVAFAVKH STVFDNLPNP EDRKNYELLC

251

GDNTRKSVDD YQECYLAMVP SHAVVARTVG GKEDVIWELL NHAQEHFGKD

301

KPDNFPQLFQS PHGKDLLFKD SADGFLKIPS KMFELYLYGY EYVTALQNLR

351

ESKPPDSSKD ECMVKWCAIG HQERTKCDRW SGFSGGAIEC ETAENTEECI

401

AKIMKGEADA MSLDGGYLYI AGKCGLPVVL AENYKTEGES CKNTPEKGYL

451

AVAVVKTSDA NINWNNLKDK KSCHTAVDRT AGWNIPMGLL YSKINNCKFD

501

EFFSAGCAPG SPRNSSLAL CIGSEKGTGK ECVPSNERY YGYTGAFRCL

551

VEKGDVAFAVK DQTVIQNTDG NNNEAWAKNL KKENFEVLCK DGTRKEVTD

601

ENCHLARGPM HAVVSRKDKA TCVEKILNKQ QDDFGKSVTD CTSNFCFLQS

651

NSKDLLFRDD TKCLASIACK TYDSYLGDDY VRAMTNLRQC STSKLLEACT

701

FHKP

Color Code: >=30 >=10 >0%

Click on the “Protein View” tab and then on the “Coverage” tab to see where the SPIDER results overlap with the protein sequence. Note that SPIDER matches are shown in red while matches found by both PEAKS protein ID and SPIDER are shown in magenta:

4. Load Data

4.1 Data Formats

Before loading data files into PEAKS, you must make sure that the data is in an accessible format. PEAKS handles data files in the following formats:

- .ANZ – the zip compressed XML based file format associated with PEAKS 4.5 and earlier
- .BAF, .YEP and folders of .FID files from Bruker instruments
- .D files from Agilent QTOF instruments
- .DAT files created by BSI's ABI converter software
- .DTA: file format associated with SEQUEST software
- .MGF: file format associated with Mascot software
- .PKL: file format associated with MassLynx software
- .RAW files from Thermo Electron instruments
- .RAW files from Waters QTOF instruments
- .WIFF files from ABI/Sciex QSTAR and QTRAP instruments
- .XML format files using the mzData schema
- .XML format files using the mzXML schema
- PEAKS 5 project directories

4.2 Data Conversion

It is best to import RAW data directly, so that PEAKS can access the complete, unprocessed experimental data including the MS survey scan and retention time information. This will ensure that the PEAKS analysis does not suffer from poor preprocessing.

In order to load RAW data from different vendors, PEAKS may require third-party software to be installed. Please consult the following instructions for third-party software requirements.

4.2.1 PEAKS Project Converter

Projects created using PEAKS 5.0 are different than those produced with 5.1. PEAKS provides a straight-forward project conversion tool that will allow users to quickly enhance existing projects to be supported in PEAKS 5.1. Users comfortable with command line, will find the tool very easy. This section will cater to those unfamiliar with command line, but who wish to convert their necessary projects promptly.

There are two main methods that can be used to generate a PEAKS 5.1 project from an existing 5.0 project.

1. The first method converts a 5.0 project and stores the 5.1 project in the same directory as the original file, with the addition of “51” to the end of the file. Just enter ProjectConverter followed by the PEAKS 5.0 project location and name. For example, if the project is found in your C:\PeaksStudio5.0\derbyServer\serverDB\My5.0Project folder, enter

```
ProjectConverter "C:\PeaksStudio5.0\derbyServer\serverDB\My5.0Project"
```

This will produce a new 5.1 project in this same directory, with the name My5.0Project.51.

The quotation marks (“”) surrounding the location are not necessary in this case, but are necessary if your path contains any spaces, such would be found in use of “C:/My Documents/Projects...”.

2. To produce a new file in a different directory, the only change is the addition of leaving a blank space, and then entering the output directory, such as “C:\PeaksStudio5.1\derbyServer\serverDB\My5.1Project”

Thus the line will look like

```
ProjectConverter "C:\PeaksStudio5.0\derbyServer\serverDB\My5.0Project" "C:\PeaksStudio5.1\derbyServer\serverDB\My5.1Project"
```

If the new directory does not already exist, the program will create the directory for you. If the directory already exists, entering a dash and the letter ‘-f’ before the original project will force the ProjectConverter to overwrite any existing projects in this location.

For example:

```
ProjectConverter -f "C:\PeaksStudio5.0\derbyServer\serverDB\My5.0Project"
```

Or

```
ProjectConverter -f "C:\PeaksStudio5.0\derbyServer\serverDB\My5.0Project" "C:\PeaksStudio5.1\derbyServer\serverDB\My5.1Project"
```

You are now ready to call up these projects into PEAKS 5.1.

4.2.2 Agilent Data

PEAKS 5.1 can load native data from Agilent QTOF, provided that the MassHunter software is installed on the same computer.

4.2.3 Applied Biosystems (WIFF)Data

WIFF data from Applied Biosystems/Sciex QSTAR (or QTRAP) mass spectrometers can be loaded, provided that the Analyst QS (Analyst 1.4.1 for QTRAP) software and the MSX plug-in are installed on the same computer as PEAKS 5.1. The MSX tool is produced and sold by Infocromics Ltd., and is available (at cost) from Bioinformatics Solutions Inc. Please contact a BSI sales representative to obtain an evaluation or full license.

4.2.4 Applied Biosystems (4700/4800)Data

BSI has created a converter to extract the data from an ABI-Oracle database. If you require this separate, free tool, contact your sales representative. Once installed, you can start up the ABI 4700 Data Extractor from the Start menu.

4.2.5 Bruker Data

PEAKS 5.1 can load data from Bruker mass spectrometers provided that the CompassXport software is installed on the same computer. If loading .fid files, which are stored in a network of folders, select the top level folder to load them all at once.

4.2.6 Shimadzu Data

Shimadzu mass spectrometer data can be loaded, provided that the Shimadzu software is installed on the same computer as PEAKS 5.1.

System Requirements

This extractor can be installed on the same machine as ABI 4700 Explorer and the Oracle database (we will call this machine the 4700 SERVER in the following instructions) or another machine that has direct network access (we require that there is no firewall or proxy) to the 4700 SERVER. Windows 2000 or Windows XP is recommended for use with this tool.

Configuration

Before using the ABI 4700 Data Extractor, it must be configured. To do so, choose “Settings” from the “File” menu. Configuration requires the following:

4700 SERVER Name or IP Address: input “localhost” if the Extractor is running on the 4700 SERVER (this is the default value), otherwise enter the IP address of the 4700 SERVER.

The socket used by the 4700 SERVER: this is the port that the Oracle database listens to (the default is 1521).

Username to access the Oracle database: most likely we do not need to change this (the default is “tsquared”).

Password to access the Oracle database: mostly likely you do not need to change this one either.

Data extraction procedure

1. Load Spot Set List from the database: (Do it via menu File | Load Spot Set List). The extractor will export the peak list of a spot set into a PKL file.

2. Open a Spot Set: (menu File | Open Spot Set) Spot Set Chooser will help the user to choose a spot set. After selecting a spot set, click 'OK' to open it. The job run information of a spot set will be shown.
3. Select a job run: There is a button to select before each job run. Only the MS/MS job run can be selected for export, as the precursor information is needed. Select a job run and click 'Convert' to do the extraction.
4. Choose a filename to save: After clicking the 'Convert' button, the user needs to input a file name and the peak lists of the selected job run will be exported.

4.2.7 Thermo Data

RAW data from Thermo Electron mass spectrometers can be loaded, provided that the XCalibur software is installed on the same computer as PEAKS 5.1.

4.2.8 Varian Data

A conversion tool is embedded into Varian's data acquisition software which allows the conversion of Varian raw data into .pkl files which can be immediately read by PEAKS.

The .trans type data (raw) is converted in Varian programs by clicking "File", "Save As" and selecting the .pkl file format or by clicking "File", right clicking "Export" and selecting ".pkl". If you are viewing a chromatogram with the Varian software, all the spectra data in the viewed chromatogram is converted to the .pkl format. Likewise, if you are viewing a single spectrum and choose to convert the data, only the viewed spectra will be converted.

4.2.9 Waters/Micromass (MassLynx) Data

PEAKS 5.1 can import RAW data from Waters/MicroMass QTOF instruments using a utility called wolf.exe (originally created as part of the Sashimi Project) to access MassLynx libraries and convert the data. PEAKS provides a version of wolf.exe compatible with MassLynx 4.1. If you need a different version of wolf.exe, please visit:


www.bioinfor.com/products/peaks/support/watersmicromass.php

Additionally, you must make sure that the following MassLynx libraries are installed on the same computer as PEAKS and wolf.exe:

- DACServer.dll
- Genutil.dll
- MetaGD32.dll
- raw.dll
- securityAccess.dll
- securitySettings.dll
- securitySignature.dll

4.3 Create a New Project

- 1) To create a new project, select “New Project” from the file menu or using the “New project”

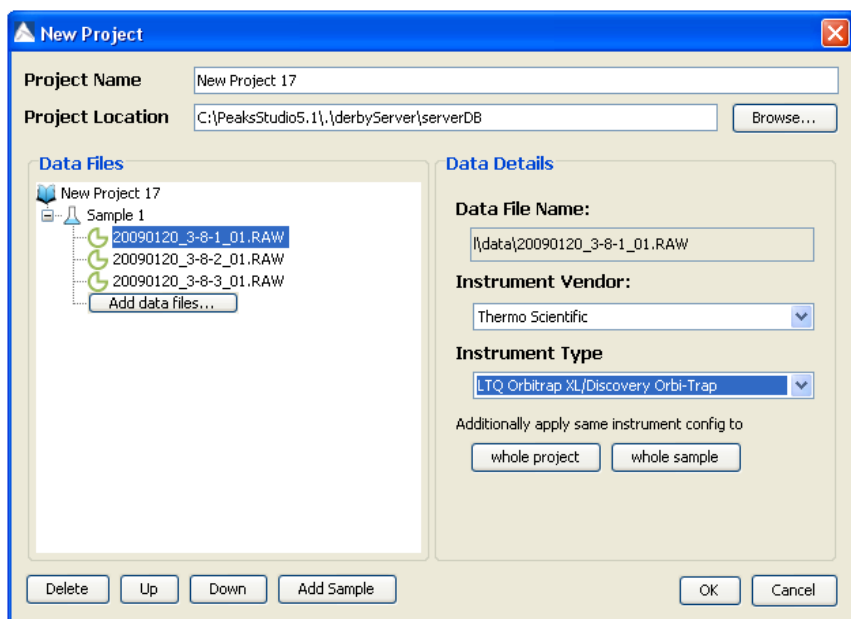
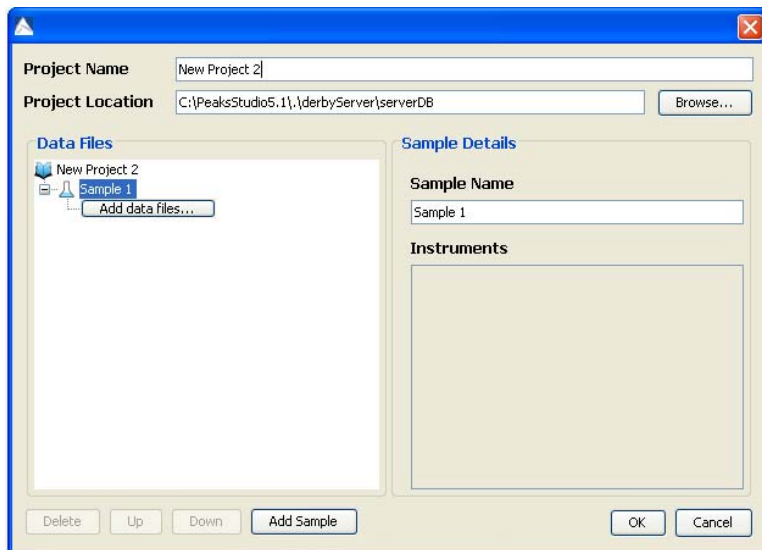
icon  on the toolbar. The following window will open:

- 2) Create a name for your project and click “Browse” to select the where you would like to save the project.

- 3) To add data files to sample 1, click the “Add data files” button. Browse for the data file and click the “Open” button. Once the data file appears, select the

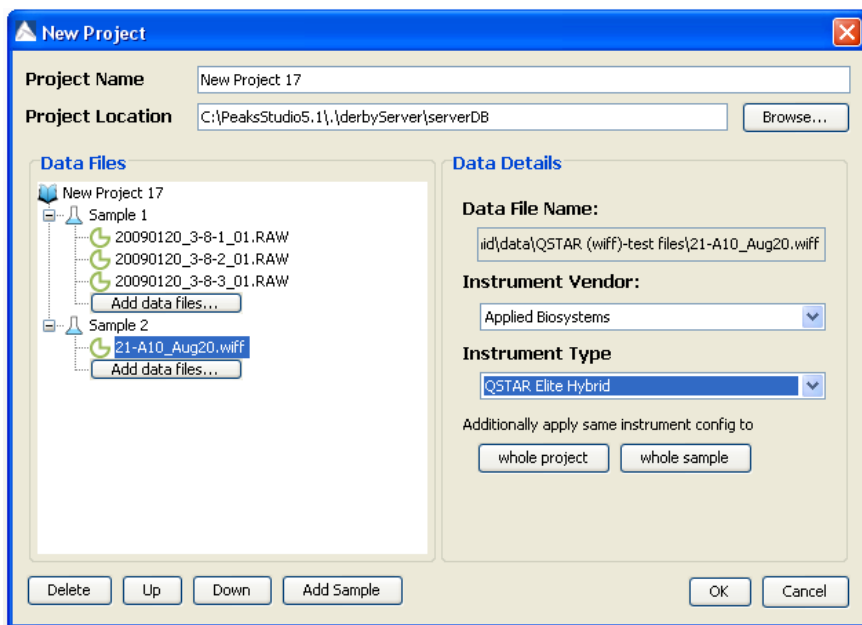
“Instrument Vendor” and “Instrument type” that was used to generate the experimental data from the drop-down lists. Select “All Instruments” from the “Instrument Vendor” drop-down list to view general instrument types in the “Instrument Type” drop-down list. If you would like to apply the same instrument configuration to all of the files in the sample, or to all of the samples in the project click on the “whole sample” button or whole project button, respectively.

- 4) To add another sample, click on the “Add Sample” button. To add a data file to sample 2 click on the “Add data files” button. Select the instrument vendor and type from the drop-down menus unless you had previously applied the instrument configuration to the whole project in step 3.



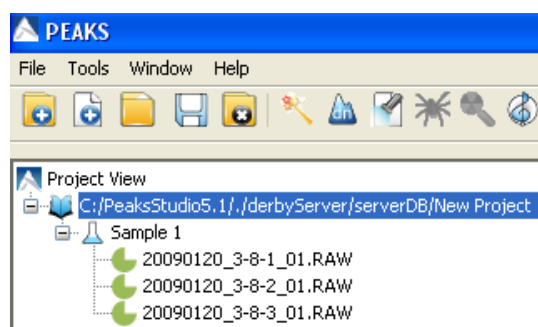
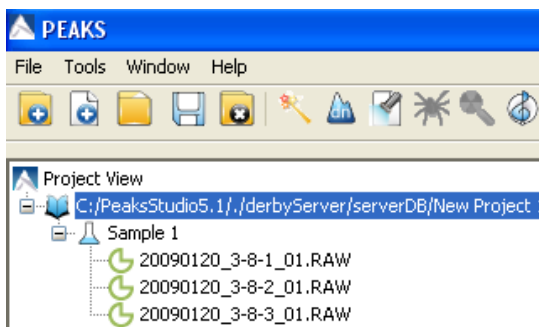
- 5) To delete a sample or data file, select the appropriate node (sample or data file) and click the “Delete” button. You can also change the order of the samples within a project or data files within a sample using the “Up” and “Down” buttons.

- 6) Click the “OK” button once all data files and samples are added to the project.




- 7) The project will appear in the “Project View” window. The “G” symbol indicates that the file is still loading (see example below).



The “G” symbol indicates that the file has finished loading (see example below).



4.4 Add Data to an Existing Project

- 1) Select “Add data” from the file menu or “” from the task bar. The original project window will open. You can add more files to an existing sample or create additional samples. You will need to select the instrument vendor type. For more information on adding files/samples or setting up the instrument configuration, refer to section [4.3 Create a New Project](#).

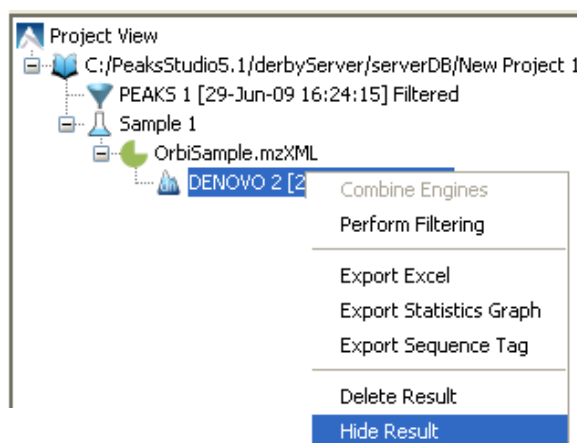
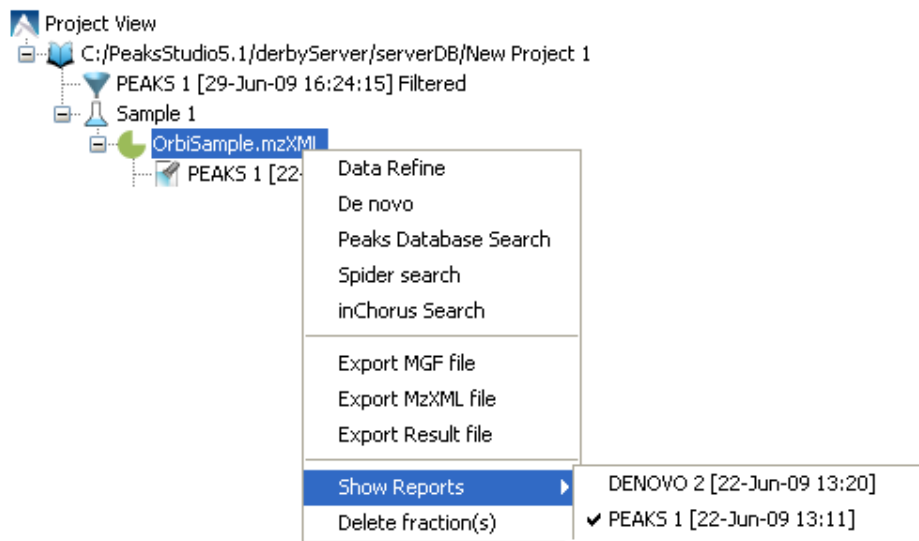
4.5 Open / Close a Project

To open a stored project, select “Open Project” or “Open Recent Project” from the file menu or “” from the task bar. To close a project that you are working on, select “Close Project” from the file menu or “” from the task bar.

4.6 Delete and Hide Results


Results and filtered results can be hidden or deleted by clicking right clicking on a results file:

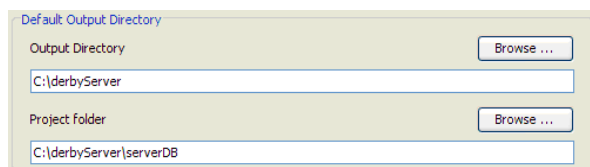
Hidden results can be brought back by right clicking on the data file and selecting “Show Reports”. Check the results that you would like to view.



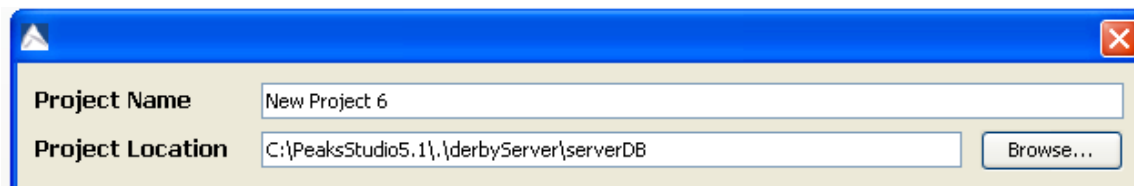
You can delete results, filtered result, fractions, and samples. Note that deleting the latter two will delete any results that depend on them as well. This can be the case even if they only depend on those results partially. While hiding results can be reversed, deleting is permanent.

4.7 Change the Location of Saved Projects

Projects are saved in the location that is listed in your “Preferences” window. To modify your preferences, select the “Preferences” toolbar icon  or select “Preferences” from the “Window” menu. Select “General” on the left hand side of the window. The default “Output Directory” and “Project Folder” locations are listed in the “Default Output Directory” panel. Please note that the defaults seen here may differ from your default locations depending where you downloaded PEAKS. Click on the “Browse” buttons to change either of these locations.



You can also change the location of your projects on a project by project basis by selecting a new “Project Location” when setting up a new project as seen below:



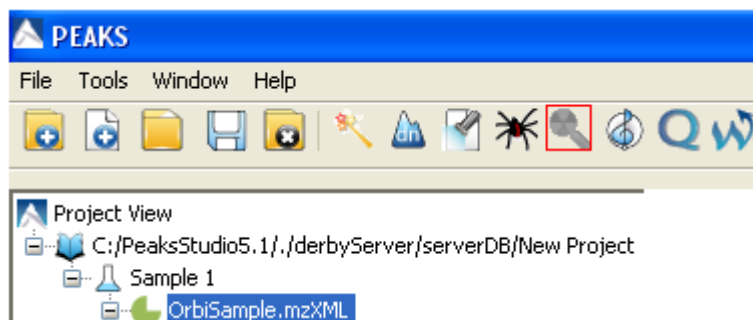
4.8 Orienting Yourself

4.8.1 Project View Panel

This frame appears in the upper left hand corner, displays the organization of a particular project (if applicable) or simply of a data file. Use the ‘+’ and ‘-’ boxes to expand and collapse the project in order to access the data file that you want to analyze. Make sure the data file to be analyzed is selected. If you select a project, all of the samples will be analyzed. If you select a sample, all of the files within that sample will be analyzed.

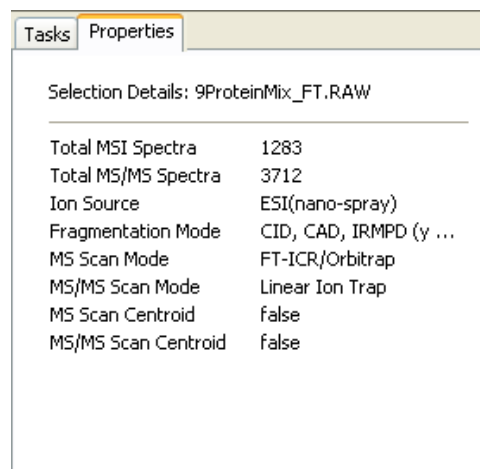


Note: Any features that cannot be performed when a specific node is selected will be grayed out. In the example below, you cannot run the PTM finder when the raw file is highlighted; you must have a PEAKS Protein ID selected in order to run the PTM finder.



4.8.2 Properties Panel

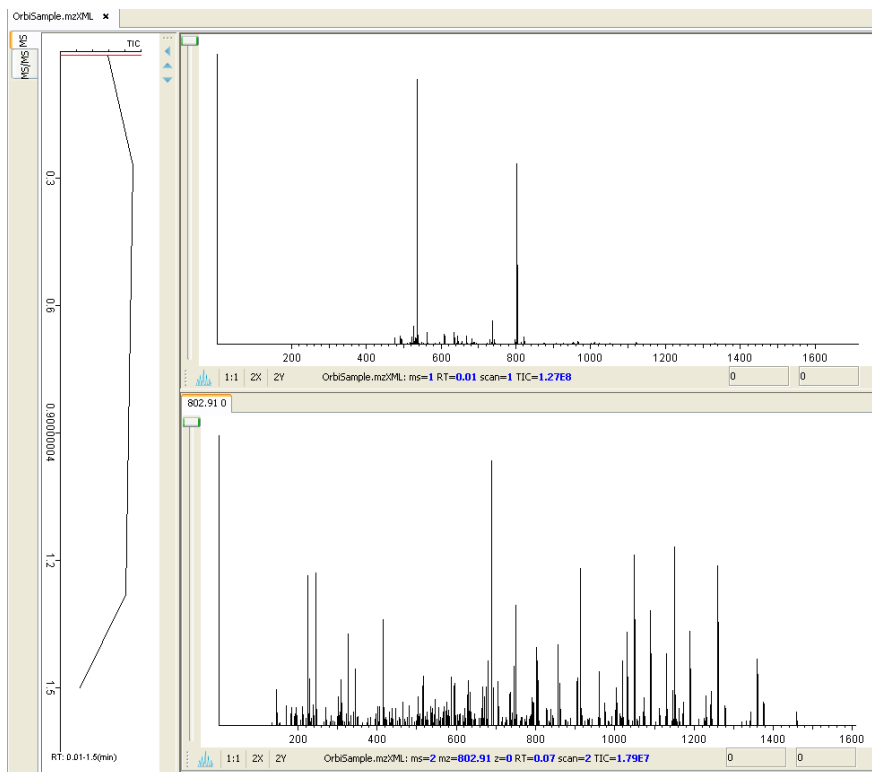
PEAKS reads and tracks information about the experiment for use in the analysis and for future reference. Once the data file has loaded, click on the properties tab in the bottom left hand corner. If any information cannot be found in the file, PEAKS will prompt you to enter this information.



4.8.3 Raw Spectrum View

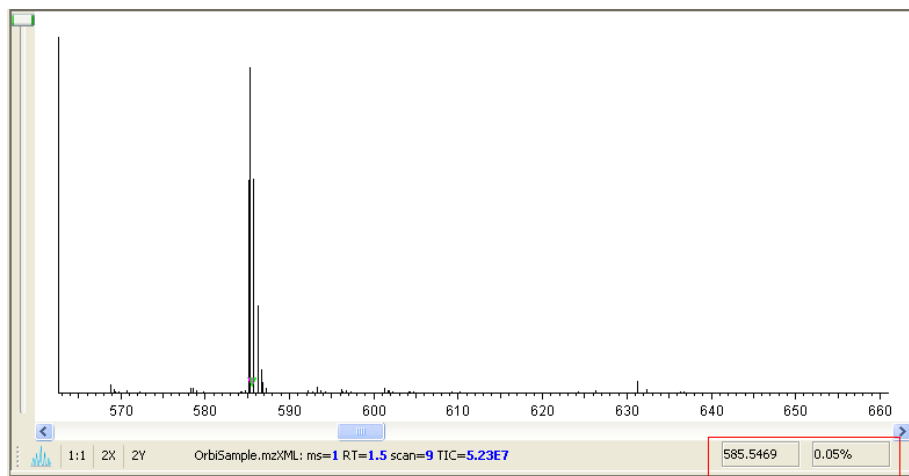
Opening the raw file in PEAKS will display the following graphs in the “Main Processing Window”.



The “MS” tab is selected by default and represents the precursor scan. On the left hand side of the screen is the total ion current (TIC). Depending on how the file was generated there may be simply a list of spectra and not a TIC graph. The retention time is plotted against the vertical axis. Clicking on the TIC graph will move the red line and display the ms spectra to the right of the TIC graph that corresponds to the selected retention time. Alternatively, use the up and down arrows, found on the keyboard, to move through the TIC. If the MS2 scans is available, it will be displayed below the corresponding ms scan.



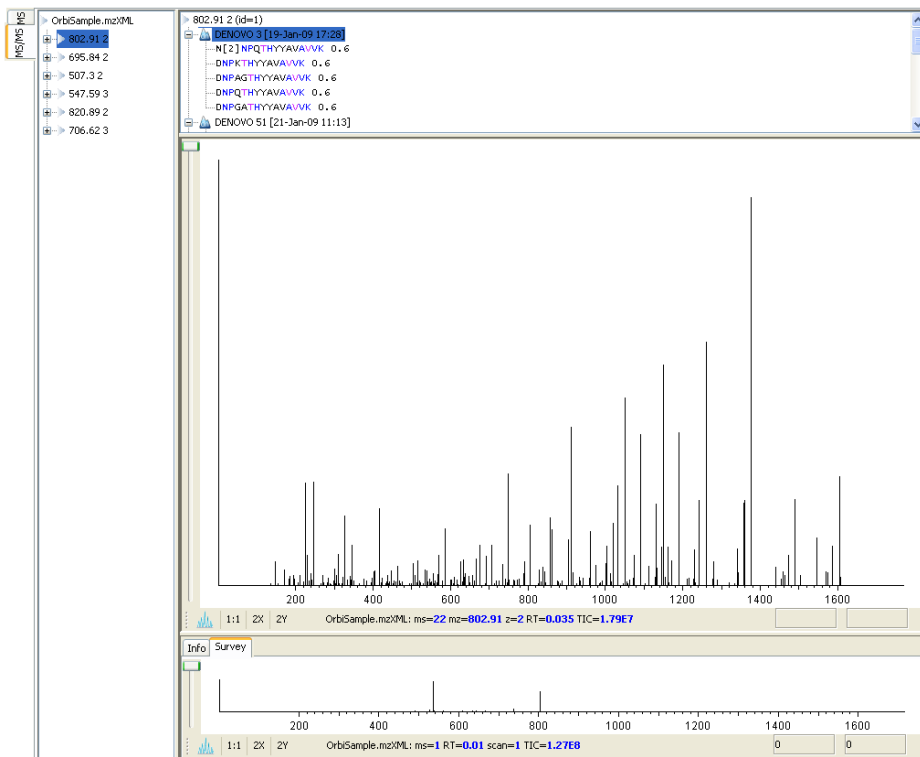
To zoom either on the X or Y axes, select the “2X” or “2Y” buttons, respectively. To scroll in even more, click the button on the left of your mouse and drag the arrow to the side. To increase the intensity of the peaks, use slide the scroll bar on the left hand side, up and down. Selecting the “1:1” button will bring you back to the original image where the entire spectrum is visible.

Scrolling over the spectrum will display the m/z ratio and the height/intensity (as a percentage of 100) of the particular peak under the spectrum view on the right hand side (see the box highlighted in red below).

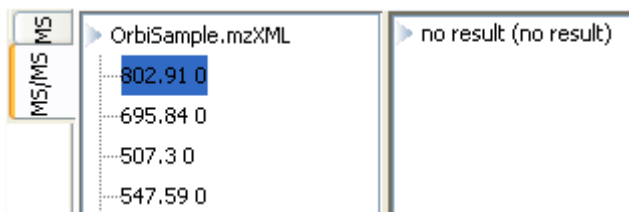


You can use the profile  and peak  buttons to switch the spectrum view from profile mode to peak mode and vice versa.

The MS/MS tab gives detailed information about each tandem spectrum:



Each of the spectra in the data file will be listed in the left most panel under the name of the data file. Clicking on one of the spectra will display the results that have been generated for that spectrum in the top right hand panel (as seen in the example above). Before any results files have been generated, the top panel will look like this:



Info Survey

Selected MS/MS [OrbiSample.mzXML, 802.91 2]
 Retention Time: 0.035
 TIC: 1.79E7
 Number of Peaks: 222
 Fragmentation Type: Unavailable

Number of Results: 7
 DENOVO 3 [19-Jan-09 17:28] has 5 matches
 DENOVO 51 [21-Jan-09 11:13] has 1 matches
 PEAKS 1 [19-Jan-09 17:22] has 1 matches
 PEAKS 2 [19-Jan-09 17:26] has 1 matches
 PEAKS 3 [19-Jan-09 17:29] has 1 matches
 DENOVO 1 [19-Jan-09 17:21] has 5 matches
 DENOVO 2 [19-Jan-09 17:25] has 5 matches

More information can be found about the spectra under the “Info” tab. You will find information about the retention time, where to find the spectra on the TIC graph, the number of peaks and the fragmentation type (if available). You will also find an overview of the results that were found for that spectrum in the results files.

The largest panel displays the MS/MS and below you will find the corresponding MS spectra under the “Survey” tab. Information about navigating through the MS and MS/MS spectra can be found above in the section describing the “MS” tab.


5. Data Refinement

Since mass spectrometry data most often contains noise and redundant data, it makes sense to filter the data before analysis. This will increase the quality of the results, while saving time spent on database searching and *de novo* sequencing.

When PEAKS Studio or Client are connected to a PEAKS Online server, you will also save time by uploading smaller, preprocessed data. Data refinement can be done locally, before uploading to the server.

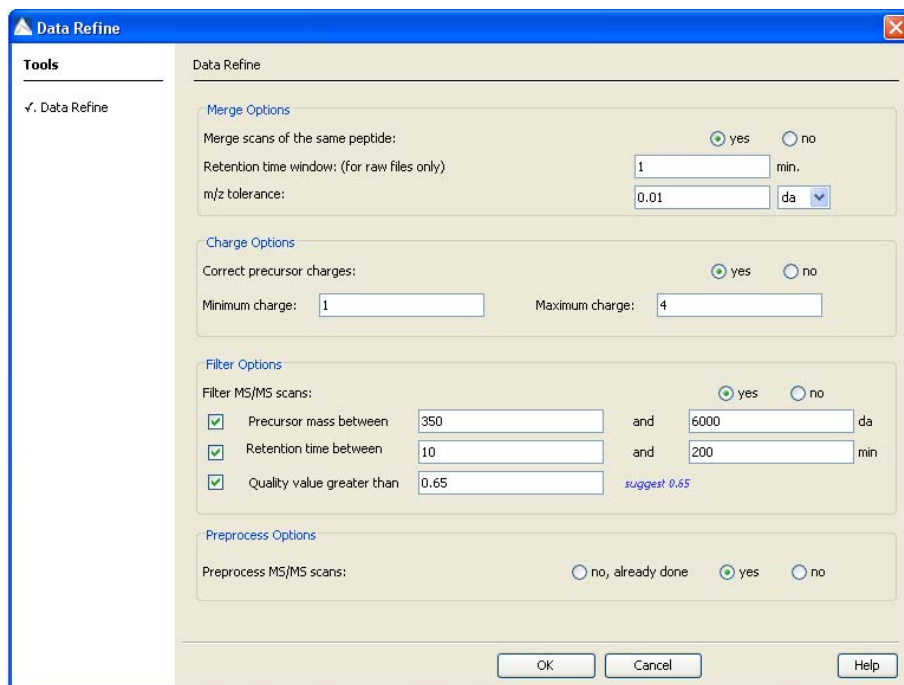
5.1 Run Data Refine

To begin the refinement of data from a whole MS/MS run:

- 1) In the “Project View Frame”, select the data file(s) containing the data that you wish to refine.
- 2) Click the Data Refine toolbar icon 

Or

Select “Data Refine” from the “Tools” menu.



The Data refinement options window will appear:

- 3) Choose the data refinement tools you wish to use by clicking the “yes” radio button next to each one. See the information below to help you decide on proper refinement parameters.

5.2 Data Refinement Parameters

5.2.1 Merging Scans

In DDA mode, a mass spectrometer will often produce several tandem ms (MS/MS) scans of the same peptide. To increase the intensity of real signal peaks within these scans and to reduce the size of the whole data set, it makes sense to merge MS/MS scans of the same peptide together. To avoid improper merging (of MS/MS scans of different peptides) we make sure that the measured parent ion masses of these peptides are very close and that they have similar retention times in the LC column. The units here are m/z values in Daltons or ppm. For retention time, we use whatever units are recorded in the data file (usually minutes or seconds).

5.2.2 Precursor Charge Correction

Since a mass spectrometer measures mass-to-charge ratios, we must know the charge on a peptide before we can determine its mass. The standard method of finding the charge is to look at the spacing of the isotope ladder in the survey scan. However many Ion-Trap instruments do not have enough resolution for this. So PEAKS will look at the MS/MS data to determine if it's charge 1+, 2+ or 3+. For data where the survey scan is available, PEAKS will examine the precursor ion's isotope distribution to confirm or correct the charge assignment. Type in the boxes to set a range of charges. Only spectra that fit in this range will be considered for analysis.

5.2.3 Filtering MS/MS Scans

Scans of contaminants and electrical noise should not be included in analysis. Removing them from the data set will save time, and reduce the risk of random matches to the database. PEAKS offers an effective tool for removing these low quality MS/MS scans. Type in the boxes to set ranges of retention time and m/z ratio. Only peaks between these values will be considered for analysis. Additionally, PEAKS examines the MS/MS spectrum to determine its quality. The quality filter is based on four characteristics: signal to noise ratio over MS/MS, number of peaks after pre-processing, sum of all peak intensities and length of the longest simple sequence tag that can be generated. You can choose a threshold of quality score (a value from 0 to 1) for accepting a scan. We recommend a quality filter of 0.65.

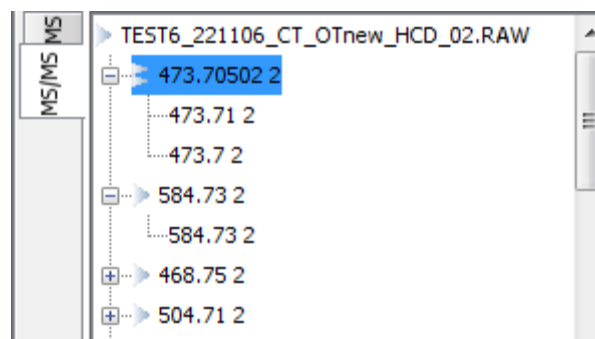
5.2.4 Preprocessing MS/MS Scans

This section deals with deconvolution (de-isotoping), centroiding and noise filtering within the MS/MS data. Preprocessing can save hard disk space or upload time. But make sure to have the original data available in case you need to refer to it later.

To see how your data is changed after data refinement, refer to the data properties window.

5.3 Data Preprocessing Results


To view the result of data pre-processing, click on the MS/MS tab on the spectrum view. In following example, the spectrum ($m/z = 473.70502$) results from raw spectra ($m/z = 473.71$ and $m/z = 473.7$).



6. *de novo* Sequencing

6.1 Setting up Auto *de novo* Sequencing Parameters

1) In the “*Project View Frame*”, select the data file(s) or project containing the spectra that you wish to sequence by auto *de novo*.

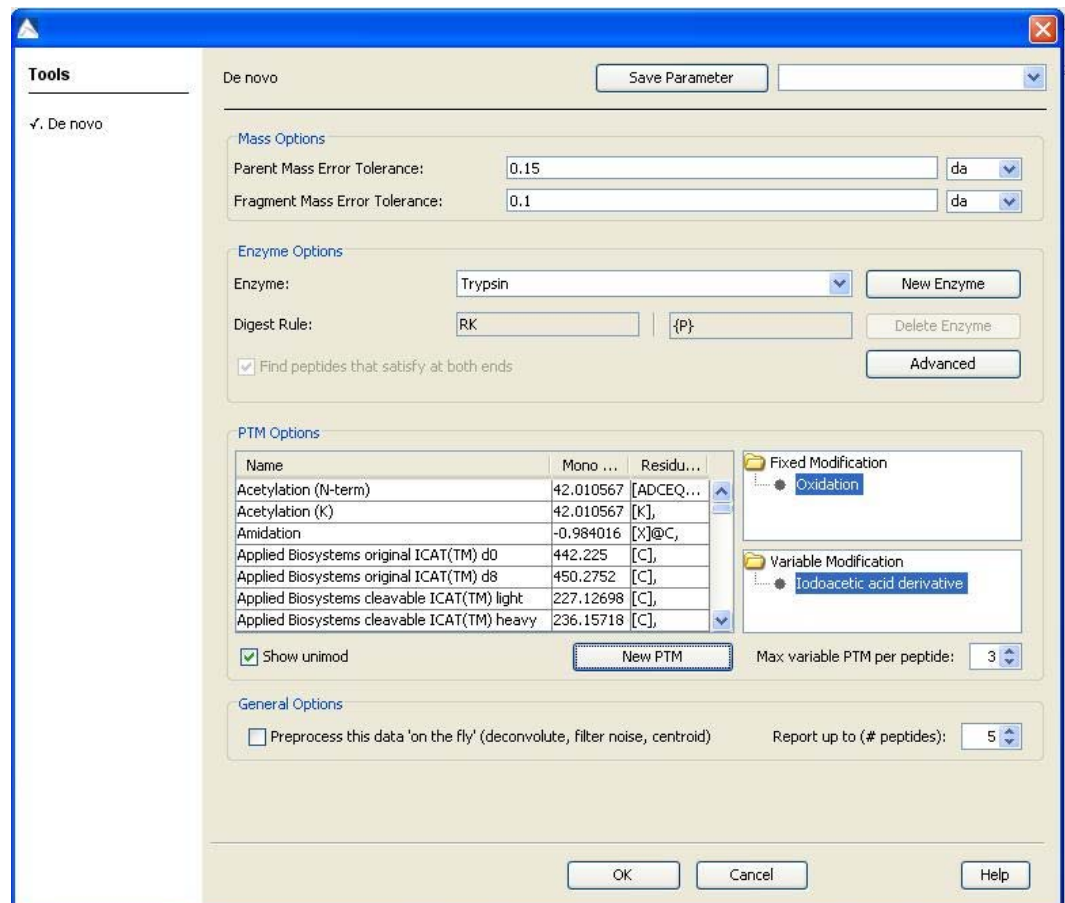
2) Click the automatic *de novo* toolbar icon 

Or

Select “auto *de novo*” from the “Tools” menu.

The auto *de novo* parameters dialogue window will appear:

3) To change any of the following parameters, now is the time:



Tools

De novo

Save Parameter

Mass Options

Parent Mass Error Tolerance: 0.15 da

Fragment Mass Error Tolerance: 0.1 da

Enzyme Options

Enzyme: Trypsin New Enzyme

Digest Rule: RK {P} Delete Enzyme

☒ Find peptides that satisfy at both ends Advanced

PTM Options

Name	Mono ...	Residu...
Acetylation (N-term)	42.010567	[ADCEQ...]
Acetylation (K)	42.010567	[K]
Amidation	-0.984016	[X]@C,
Applied Biosystems original ICAT(TM) d0	442.225	[C]
Applied Biosystems original ICAT(TM) d8	450.2752	[C]
Applied Biosystems cleavable ICAT(TM) light	227.12698	[C]
Applied Biosystems cleavable ICAT(TM) heavy	236.15718	[C]

☒ Show unimod New PTM

Fixed Modification

- Oxidation

Variable Modification

- Iodoacetic acid derivative

Max variable PTM per peptide: 3

General Options

☐ Preprocess this data 'on the fly' (deconvolute, filter noise, centroid) Report up to (# peptides): 5

OK Cancel Help

6.1.1 Mass Options

Parent mass error tolerance: Determine how much random and systematic experimental error on the parent/precursor ion PEAKS will allow for in its analysis. As you have previously selected your instrument, PEAKS will provide the suggested error tolerances. Type a tolerance in the textbox and choose units from the drop-down list. Using PPM allows for larger errors at larger m/z values. PEAKS will be very stringent concerning this value, so new PEAKS users should try setting this a little higher than past experience may suggest, if sensitivity is a concern.

Fragment mass error tolerance: Determine how much random and systematic experimental error on the fragment/daughter ion PEAKS will allow for in its analysis. As an instrument has previously been selected, PEAKS will provide suggested error tolerances. Type a tolerance in the textbox. Again, new PEAKS users should try setting this a little higher than past experience may suggest.

6.1.2 Enzyme Options

Enzyme: Tell PEAKS what type of enzyme was used to digest the sample. Choose from a drop-down list of enzymes, or if your enzyme is not in the list, click the "New Enzyme" button. You can then input the name of the new enzyme.

Digest Rules: Enter the amino acid that is found at the end of the peptide. Put set brackets {} around a residue to denote any amino acids except for those that are within the brackets. Select the "Advanced" button if your digest rules are more complicated. Select the radio box "Select peptides that satisfy at both ends" if you require that your peptide was cut by the enzyme you chose at both ends.

6.1.3 PTM Options

Selecting fixed and variable PTMs: The "PTM Options" list tells PEAKS what types of post-translational modifications to include in its analysis. To view additional modifications, select the "Show unimod" box. If a desired PTM does not appear on the list or is different than what is listed, select the "New PTM" button and the "PTM Editing" window will open. Fill in the information pertaining to your PTM. To select a PTM as Fixed or Variable, drag the PTM into the Fixed Modification or Variable Modification box. If you drag over an incorrect PTM, simply drag it back to the "PTM Options" list.

Max variable PTM per peptide: To reduce uncertainty, limit PEAKS' *de novo* sequencing 'vocabulary' by restricting the number of variable PTM found on a peptide. Specify a number by typing it into the box. To lift such restrictions, type a very large number (longer than the length of the peptide).

6.1.4 General Options

Report up to (# peptides): Set how many peptide sequences PEAKS will report from its *de novo* sequencing analysis.

Preprocess your data "on the fly" before auto de novo: PEAKS has its own built-in preprocessor for removing noise, centroiding and deconvolution. Check this box to turn preprocessing on.

BSI highly recommends using PEAKS to preprocess all data, as opposed to using instrument vendor software, if the data is to be used by PEAKS. PEAKS preprocessor should not be used on data that has already been pre-processed as this will have adverse effects on the results (unless it is ion-trap data).

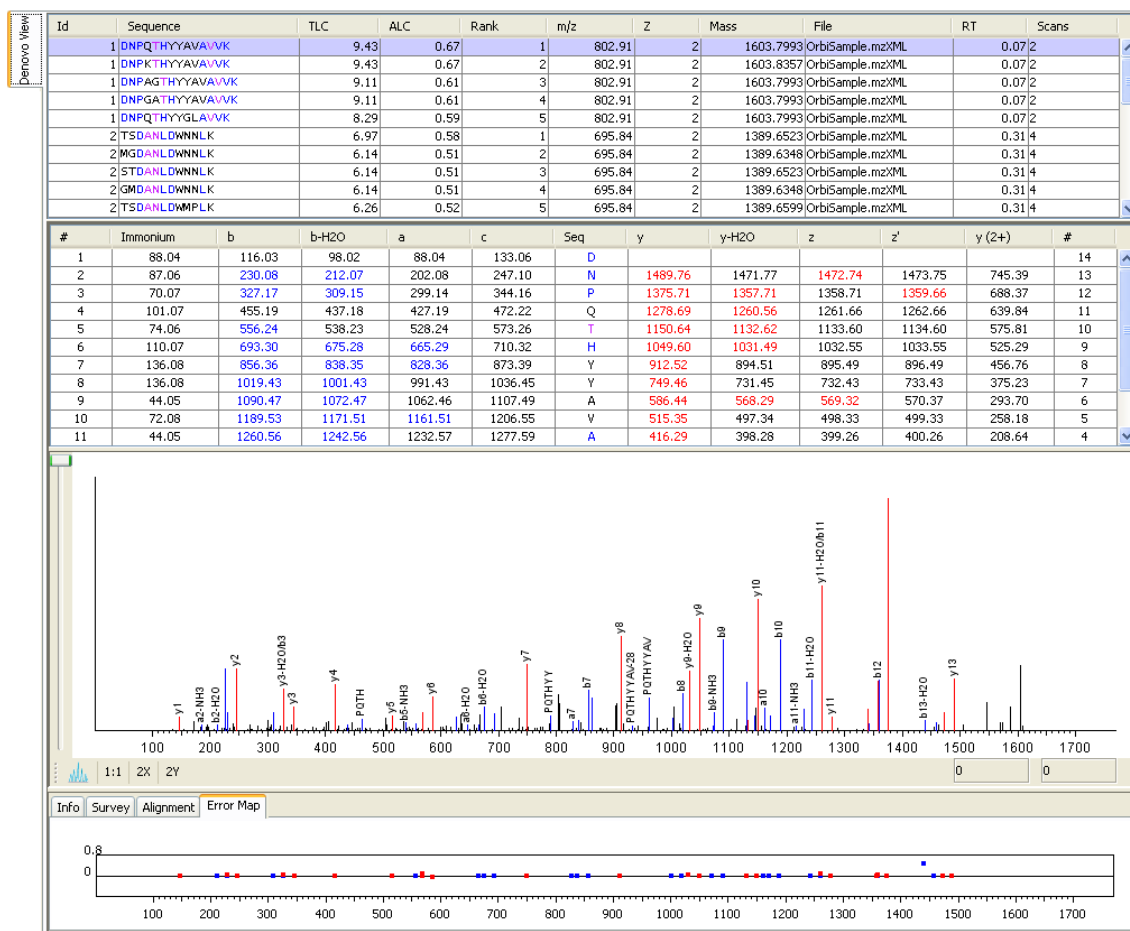
Note: If you have already pre-processed your data in the data refinement step, you do not need to do this again.

4. After setting parameters, you can save them for future use. Click the "Save Parameters" button at the top of the window, and choose a name for future reference when prompted. Any parameters that are saved will be available in the drop-down list at the top of the window. To see what's inside, select one and the parameters boxes will be populated.

5. Press the "OK" button to initiate *de novo* sequencing.

6.2 *de novo* Sequencing Results

Once *de novo* sequencing is finished, the following window will open:




6.2.1 Peptide Candidates Frame

PEAKS displays the peptide sequence candidates at the top of the screen in the “Peptide Candidates Frame”. You can sort the results by clicking on any of the titles of the columns. For example, to sort the peptide sequence candidates by ID click on “ID”. Note that all of the peptides that have the same ID have the same mass, charge, retention time and quality score (See section [5.2.3 Filtering MS/MS Scans](#) for more information on how the quality score is generated).

The following table describes the contents of the columns in the “Peptide Candidates Frame”:

ID	A unique identifier for the MS/MS spectrum. <i>This differs from a scan number since we may have merged several scans together.</i>
Sequence	The sequence of the peptide (including modifications if present) as determined by <i>de novo</i> sequencing.
TLC	Total local confidence (the confidence that we have in the peptide sequence). It is calculated by adding the positional confidence for each amino acid in the peptide sequence.
ALC	Average local confidence (the confidence that we have in the peptide sequence). It is calculated by adding the positional confidence for each amino acid in the peptide sequence and dividing by the total number of amino acids.
Rank	The sequences for a particular spectrum (ID) as sorted by score (TLC).
m/z	The measured mass/charge value, in Daltons, for the peptide.
Z	The calculated charge value for the peptide.
Mass	Calculated using the measured m/z and calculated z, we use this as the experimental mass of the peptide.
File	The name of the file.
RT	Retention time (elution time) for the peptide as recorded in the scan header.
Scan	The scan number.
Quality	A value from 0 to 1 estimated from the spectrum to refer to spectrum quality. Attributes like signal to noise, total intensity, and spectrum tagging are used.
Scan Mode	Mode that the scan step was performed in.
Frag. Mode	Mode that the fragmentation step was performed in.

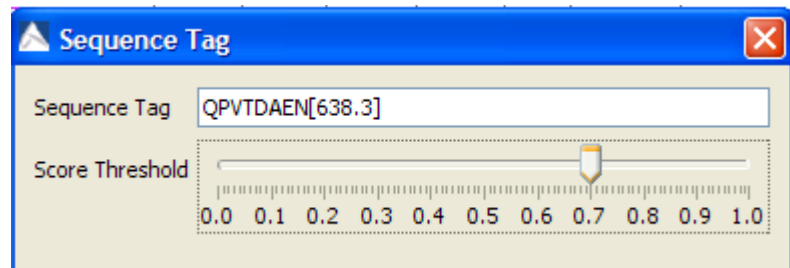
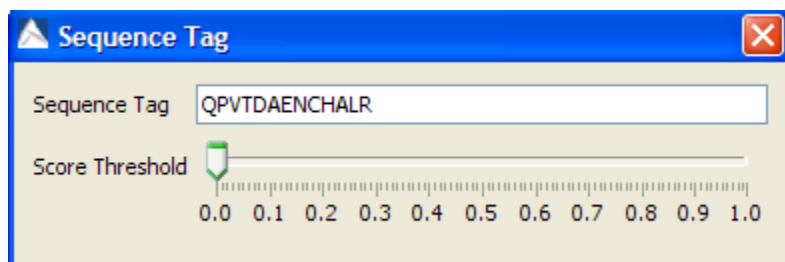
The columns themselves are customizable. To change the columns are displayed click on the “” button.

Confidence Scores

Next to the proposed sequence candidates, the *auto de novo* “Total Local Confidence” (TLC) and “Average Local Confidence” (ALC) confidence scores are shown. The confidence scores for each amino acid (that is, confidence that the correct residue in each position has been identified) are represented by *color coding*. **Red** represents a very high confidence (greater than 90%), **purple** represents a high confidence (80 to 90%) **blue** represents a medium confidence (60 to 80%) and **black** represents a low confidence (less than 60%). For more detailed positional confidence, place the cursor over the sequence of interest and right click “Show Positional Confidence”. A “Position Confidence Table” will appear, showing the confidence that each amino acid/pair of amino acids are correct.

Sequence Tags

Right click on a peptide in the “Peptide Candidates Frame” and select “Show Sequence Tag”. If the score threshold is set at 0.0, all of the amino acids in the peptide sequence will be displayed.



Increasing the “Score Threshold” will display a mass in square brackets if the amino acids do not satisfy the score threshold.

Modifications

Consider the following sequence: **DW[1]C[1]SFTDAENVQALAR**

The number 1 in square brackets refers to where a modification may occur. If you forget what modifications you selected before running *de novo*, click to the “Properties” tab.

The fixed modification is set to [1]58.005478@[CKW]. In the sequence above, the modification has been made to the W as well as the C. The *colors* assigned to the [1] follow the same confidence scores as the amino acids themselves.

Refer to the above section on “Confidence Scores” for more information on *color coding*.

Denovo details: Test 6

Parent Mass Error Tolerance	0.1 da
Fragment Mass Error Toler...	0.6 da
Enzyme	Trypsin
Semi (enzyme)	false
Fixed Modification	[1]58.005478@[CKW],
Max variable PTM per peptide	2
Report # peptides	5
Preprocess data	false

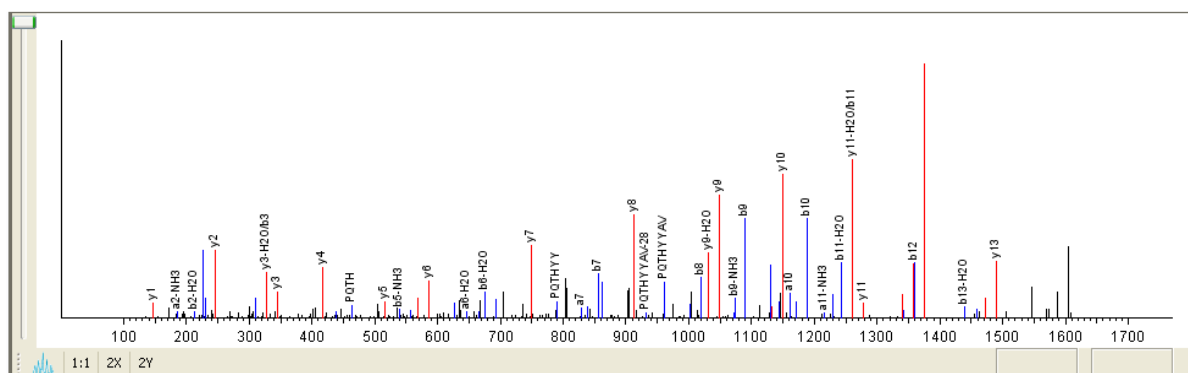
6.2.2 Ion Table Frame

#	Immonium	b	b-H ₂ O	a	c	Seq	y	y-H ₂ O	z	z'	y (2+)	#
1	88.04	116.03	98.02	88.04	133.06	D						14
2	87.06	230.08	212.07	202.08	247.10	N	1489.76	1471.77	1472.74	1473.75	745.39	13
3	70.07	327.17	309.15	299.14	344.16	P	1375.71	1357.71	1358.71	1359.66	688.37	12
4	101.07	455.19	437.18	427.19	472.22	Q	1278.69	1260.56	1261.66	1262.66	639.84	11
5	74.06	556.24	538.23	528.24	573.26	T	1150.64	1132.62	1133.60	1134.60	575.81	10
6	110.07	693.30	675.28	665.29	710.32	H	1049.60	1031.49	1032.55	1033.55	525.29	9
7	136.08	856.36	838.35	828.36	873.39	Y	912.52	894.51	895.49	896.49	456.76	8
8	136.08	1019.43	1001.43	991.43	1036.45	Y	749.46	731.45	732.43	733.43	375.23	7
9	44.05	1090.47	1072.47	1062.46	1107.49	A	586.44	568.29	569.32	570.37	293.70	6
10	72.08	1189.53	1171.51	1161.51	1206.55	V	515.35	497.34	498.33	499.33	258.18	5
11	44.05	1260.56	1242.56	1232.57	1277.59	A	416.29	398.28	399.26	400.26	208.64	4

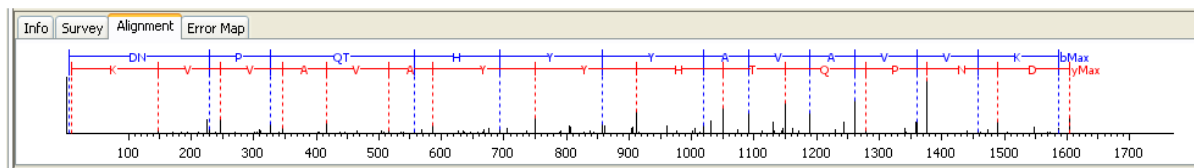
The “Ion Table” shows the proposed ions with their corresponding masses. To add additional ions to the ion table, see section [15.1.4 Ion Editor Preferences](#).

If an ion is found in the corresponding spectrum, it must first pass two criteria before being displayed in a specific *color* (blue for N-terminal ions and red for C-terminal ions). It must be found within the mass error tolerance chosen by the user and must have an intensity of greater than 2% of the ion with the greatest intensity.

6.2.3 Spectrum View Frame

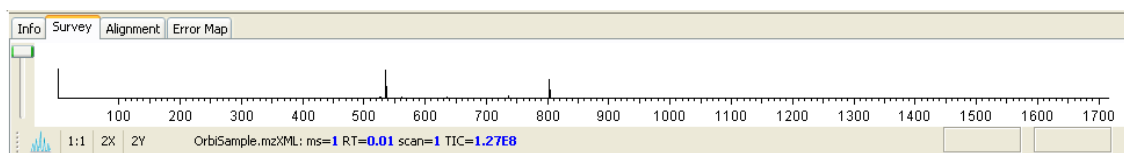


6.2.4 Spectrum Alignment Frame



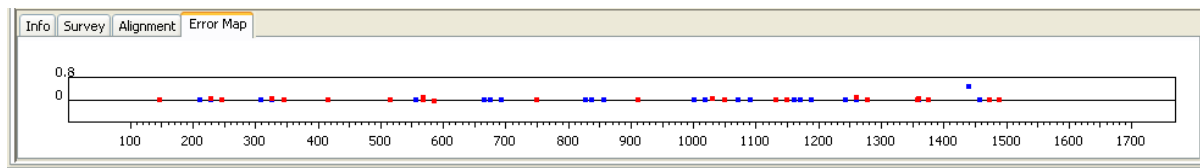
Clicking on the “Alignment” tab will display the “Spectrum Alignment Frame”. This frame always shows the whole spectrum and is used as a tool to help us navigate the spectrum view frame. A blue bar along the horizontal axis of the alignment view indicates the range of the spectrum view in the Spectrum View Frame. This frame will show you how the proposed ions align with the spectrum. By default, the Spectrum Alignment Frame displays b-ions and y-ions. The b ions are shown right to left in blue, while the y ions are shown left to right in red.

6.2.5 Survey Scan



Clicking on the “Survey” tab will display the corresponding precursor ion spectrum. The buttons that appear in this section are the same as those that are explained above in the “Spectrum View Frame” section.

6.2.6 Error Map



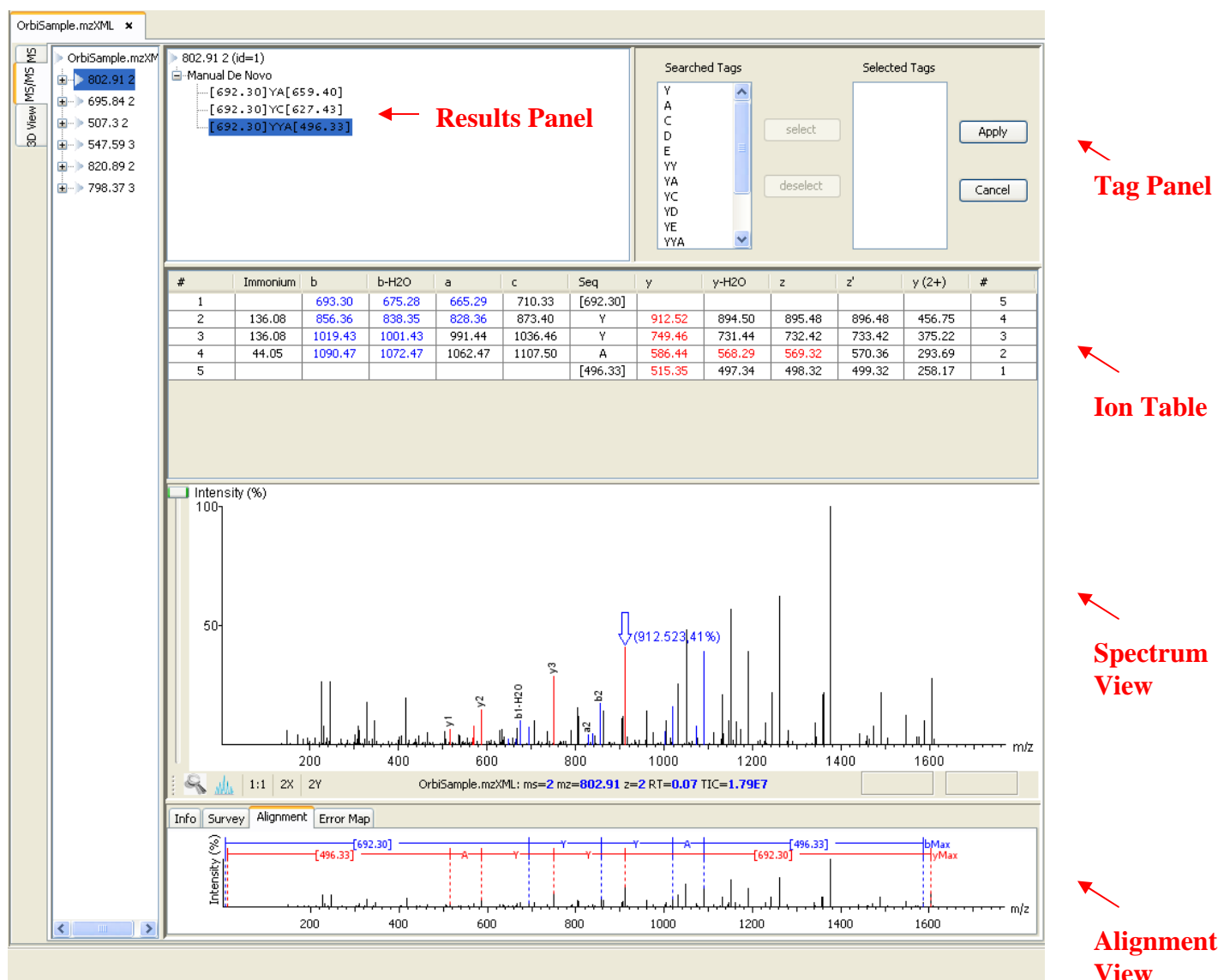
Click on the “Error Map” tab. The m/z ratio is displayed on the y axis and the error is listed on the x axis in Daltons. The “Error Plot” displays the confidence that is assigned to each ion. The most confident results lie on the centerline. Clicking a cell or column in the Ion Table highlights the corresponding points on the error plot and corresponding peaks on the spectrum.

6.3 Manual *de novo* Sequencing

PEAKS 5.1 provides a set of tools to help you sequence a peptide, using graphic cues from the spectrum.

6.3.1 Manual *de novo* Graphical User Interface

The figure below shows the main panels related to manual *de novo*. The five main panels are indicated in the figure below:



The panels are briefly described below:

The “**Result Panel**” shows all the sequencing results. The results of manual *de novo* are listed in the sub-tree with root “Manual *de novo*”.

The “**Spectrum View Panel**” shows a graphical representation of the spectrum: the peaks in the spectrum, the user-selected peaks and assigned ions. We can pick a peak on the panel and assign ions or tags to it in Manual *de novo*.

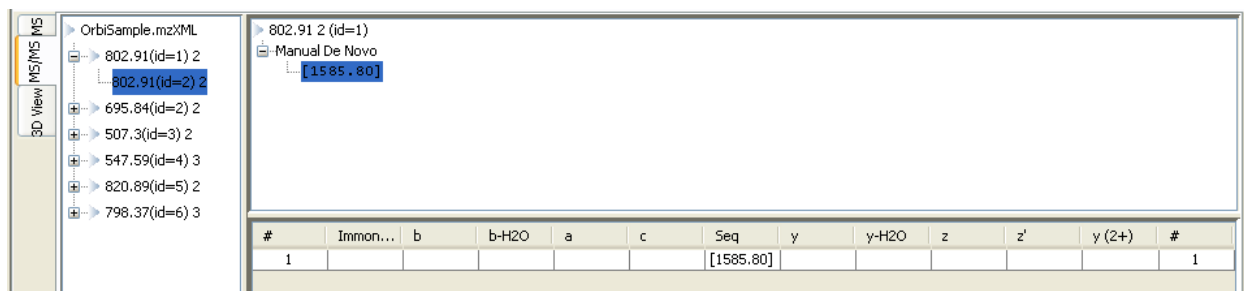
The “**Alignment View Panel**” shows how the proposed ions assigned in manual *de novo* align with the spectrum. By default, the “Spectrum Alignment Frame” displays b-ion and y-ion. The b ions are shown right to left in blue, while the y ions are shown left to right in red.

The “**Ion Table Panel**” shows the proposed ions with their corresponding masses. The default Ion Table will display immonium, b, b-H₂O, a, c, y, y-H₂O, z, z' and y(2+) ions.

The “**Tag Panel**” will appear when you search tags or ions in the spectrum. You can select the tags in the list using the “Select” button. Clicking the “Apply” button will add the selected tags to the sequence candidate.

To create a new peptide candidate for sequencing:

- 1) Right click in the “Results Panel”. This will bring up a pop-up menu.
- 2) Select “New candidate for manual *de novo*” from the pop-up menu. A new candidate will be created under the ‘Manual *de novo*’ heading. The new candidate will not have been sequenced, so it will be represented by the mass of the peptide less the mass of water (see an example below).

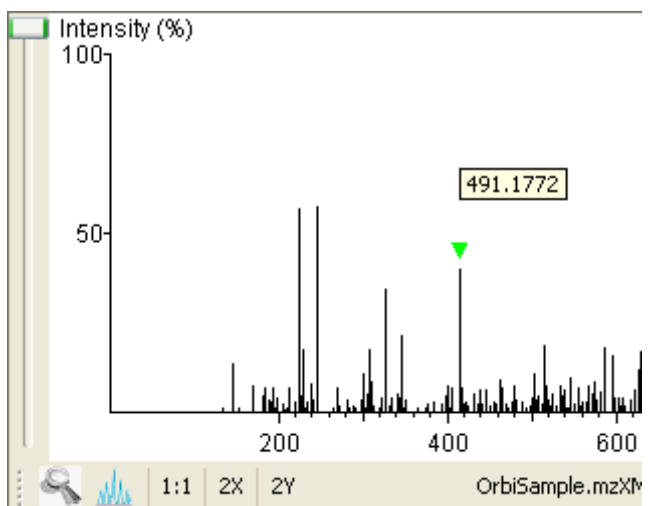


Note: The popup menu will not be accessible if you have highlighted any of the results in the “Results Panel”.

6.3.2 Manual *de novo* Operations

Note: All operations occur in the Spectrum View Frame of the Main Processing Window.

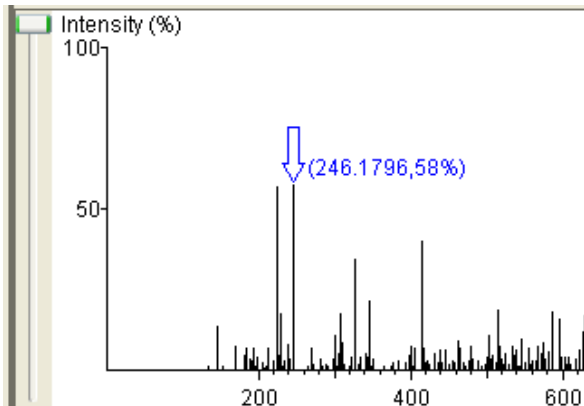
When the mouse is placed in the Spectrum View Frame, a green (by default) arrow follows the movement of the mouse. This is the *Position Bar* and it is used as a cursor for all manual *de novo* operations. The cursor's position on the m/z scale is enumerated on the top of the *Position Bar*.



Selecting a peak

To select a peak, click on it. A blue (by default) arrow, called the *Freeze Bar*, indicates the selected peak.

Alternatively an ion peak can be selected by clicking on its corresponding cell in the Ion Table.



Measure the m/z difference between two PEAKS


Select a peak (blue arrow) with the *Freeze Bar* and move the mouse to the left or right. Hold the *Position Bar* (green arrow) above another peak. The number in the *Position Bar* is the difference between the two PEAKS (in the example below the difference is 80.99109).

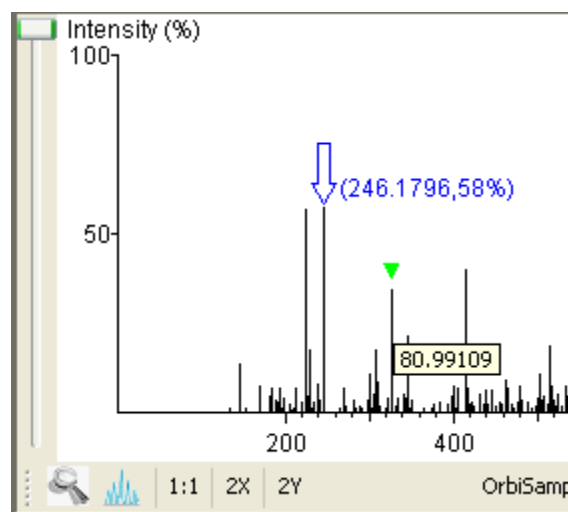
Deselect a peak

Double click anywhere in the Spectrum View Frame

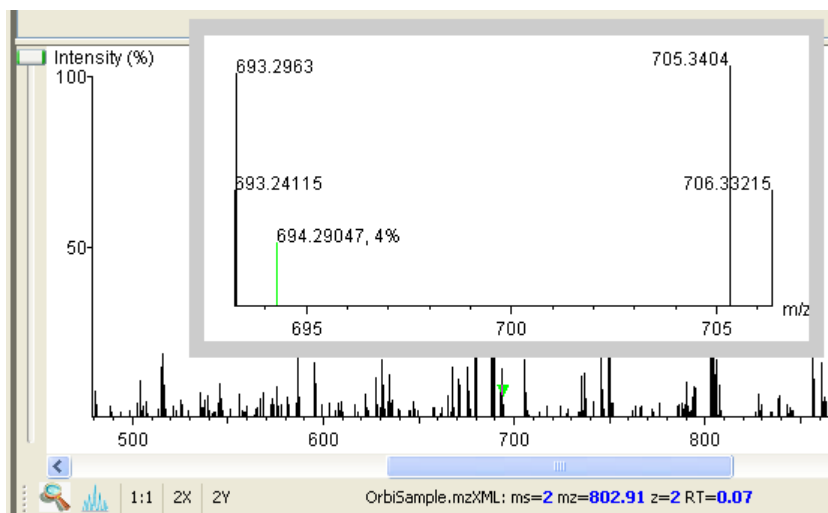
Zoom in on part of the spectrum

In the Spectrum View Frame, click and drag the mouse horizontally. The selected area will be

shown in the Spectrum View Frame. Click on the magnifying tool “” to display a panel with the spectrum zoomed in. The peak which is selected by the *Position bar* will be highlighted in green.



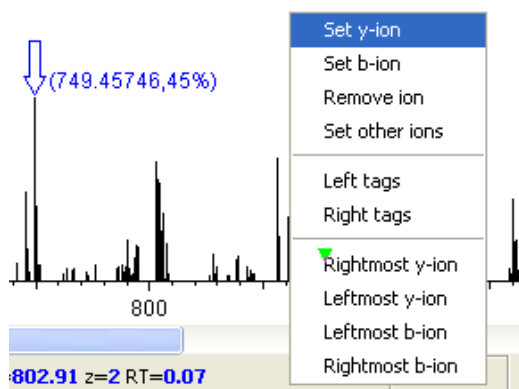
Click on the magnifying tool again to remove the zoomed panel.



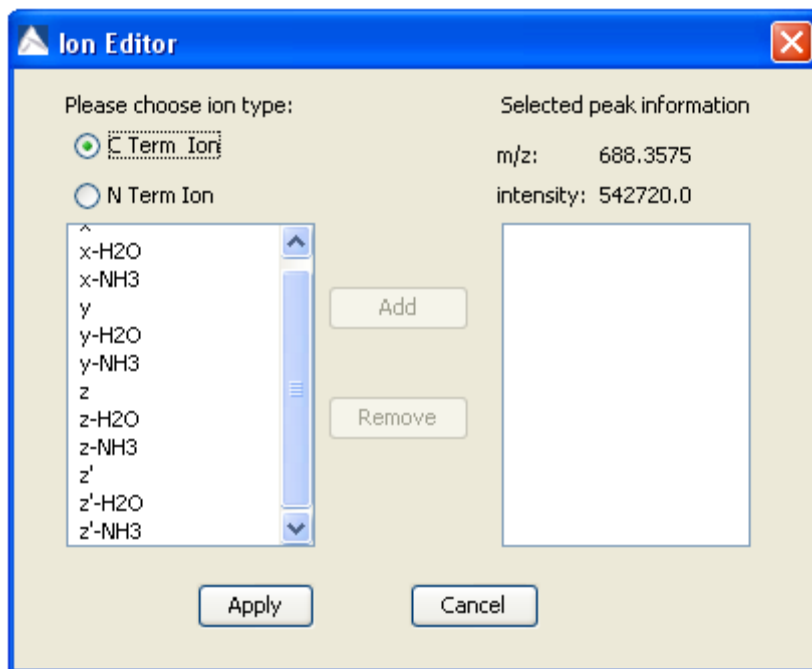
Setting/removing ions to/from a peak

Select a peak, then right click the mouse anywhere in the Spectrum View Frame. Select “Set y-ion” from the pop-up menu to designate the peak as a y ion, or “Set b-ion” from the pop-up menu to designate the peak as a b ion. Click on “Remove ion” to remove the ion that you have previously set.

Select “Set other ions” from the pop-up menu to view the “Ion Editor” dialog box. The Ion Editor dialogue allows you to add or remove ion designations to/from

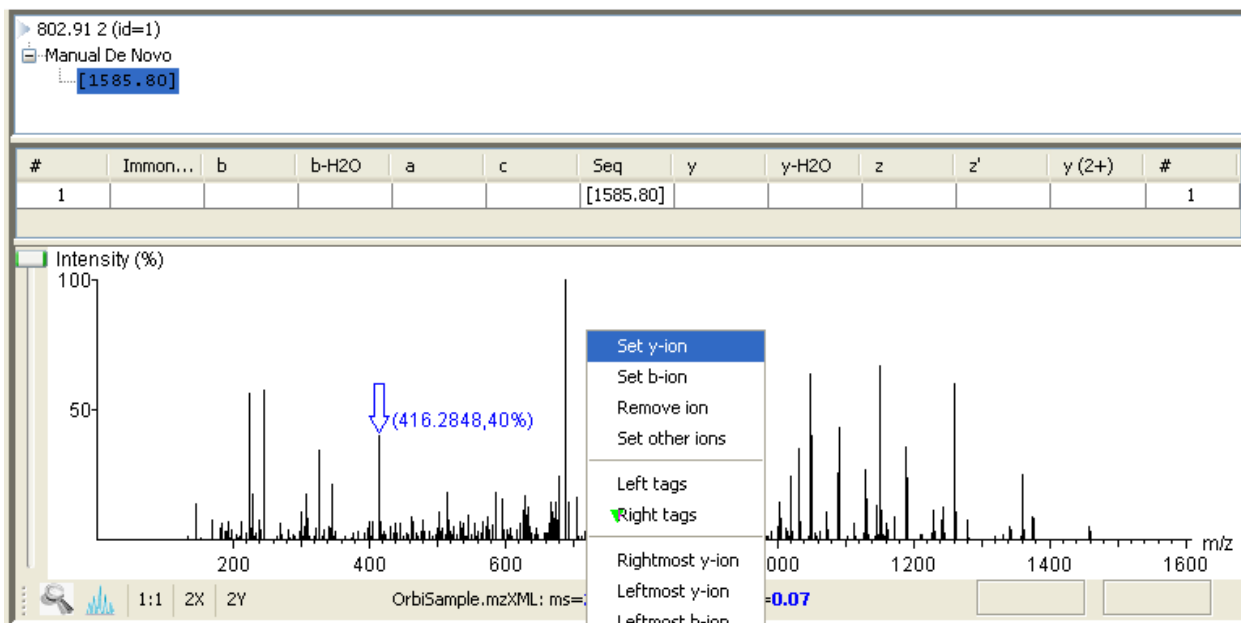


a peak. Select either “C Term Ion” or “N Term Ion” to see the C and N terminal ions respectively. Then select an ion from the ion choice list and press the “Add” button to add it to the selected ion list. Remove an ion from the selected ion list by selecting it and pressing the “Remove” button. Click “Apply” to apply the changes to the selected peak.



After setting an ion, both the alignment view and the peptide sequence candidate name (as displayed in the peptide candidate frame) will change to reflect the mass remaining to be sequenced on either side of the ion.

In the example below, the selected peak at 416.2848 m/z was designated as a y-ion.



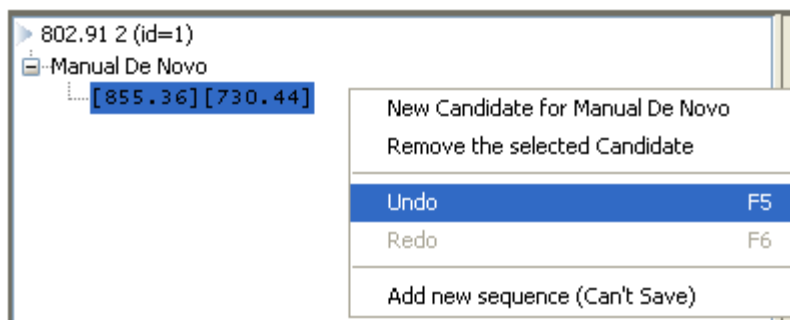
Note that the manual *de novo* candidate information is updated in the results panel, ion table and alignment view panel. The selected ions are also highlighted in the spectrum view panel.

Search a sequence tag

Select a peak with a defined ion (i.e. an ion that has been labeled with a peptide). Right click to trigger the popup menu, then select “Right tags” or “Left tags” to search peptide tags either to the right or left of the selected peak. PEAKS will select the appropriate tags and show them in the “Tag Panel” frame. You can test the suitability of a tag by highlighting it in the “Searched Tags” panel; the corresponding information for the tag will be shown in the Spectrum View panel, the Ion table and the Alignment frame panel. You can insert one or more tags by highlighting the desired tags, clicking “Select” to move them into the “Selected Tags” panel and then clicking “Apply”. Press the “Cancel” button at any time to exit the search and discard any changes.

Undoing an edit

If you have made an error in your sequencing it is possible to undo the change. With the Peptide candidate still selected in the “Results panel”, right click the mouse and select “Undo” to return to the previous peptide sequence. You can click this button multiple times to return to earlier stages in your edit.



Redoing an edit

With the Peptide candidate still selected in the “Results panel”, right click the mouse and select “Redo” if you have undone one too many changes. You can click this button multiple times to proceed to later stages in your edit.

7. Database Search

7.1 Setting up Protein Identification Parameters

1) In the “*Project View Frame*”, select the data file(s) or project containing the spectra that you wish to identify using database search.

2) Click the Protein Identification toolbar icon 

Or

Select “PEAKS Protein ID” from the “Tools” menu.

The Protein Identification Parameters dialogue window will appear:

3) To change any of the protein identification search parameters, now is the time.

Database Search Save Parameter test 6

Mass Options

Parent Mass Error Tolerance: Precursor Mass Search Type: ☒ Monoisotopic ☐ Average

Fragment Mass Error Tolerance:

General Options

☒ Preprocess this data 'on the fly'

Max Missed Cleavages:

Enzyme Options

Enzyme: New Enzyme

Digest Rule: Delete Enzyme

☒ Find peptides that satisfy the above rule at both ends Advanced

PTM Options

Name	Mono mass	Residue site
Acetylation (N-term)	42.010567	[ADCEQGILMPSTV]@N,
Acetylation (K)	42.010567	[K],
Amidation	-0.984016	[X]@C,
Applied Biosystems ori...	442.225	[C],
Applied Biosystems ori...	450.2752	[C],
Applied Biosystems cle...	227.12698	[C],
Applied Biosystems cle...	236.15718	[C],
Applied Biosystems ITR...	144.10591	[X]@N,

☐ Show unimod New PTM Max variable PTM per peptide:

Fixed Modification

- Iodoacetic acid derivative

Variable Modification

Database Options

☒ Select database Select Database: all species

☐ Paste fasta sequences New Database Edit Database Set Taxa

Advanced Options

PEAKS uses a hybrid search technique that requires some sequence tags to help in the search

☒ I have already run de novo, don't run it again

☐ Run de novo using different parameters than the above

☐ Run de novo using the same parameters as above (default)

☐ Validation - decoy search

7.1.1 Mass Options

Parent mass error tolerance: Determine how much random and systematic experimental error on the parent/precursor ion PEAKS will allow for in its analysis. Type a tolerance in the textbox and choose units from the drop-down list. Using PPM allows for larger errors at larger m/z values. PEAKS will be very stringent concerning this value, so new PEAKS users should try setting this a little higher than past experience may suggest, if sensitivity is a concern.

Fragment mass error tolerance: Determine how much random and systematic experimental error on the fragment/daughter ion PEAKS will allow for in its analysis. Type a tolerance in the textbox. Again, new PEAKS users should try setting this a little higher than past experience may suggest.

Precursor mass search type: If the precursor mass is monoisotopic value, check monoisotopic. Check average, otherwise.

7.1.2 Enzyme Options

Enzyme: Indicate which type of enzyme was used to digest the sample. Choose from a drop-down list of enzymes. Note that you cannot delete or change the details of a built-in enzyme and therefore the “Delete enzyme” button and the “Digest Rules” panel will be grayed out. If your enzyme (or combination of enzymes) is not in the list, click the “New Enzymes” button. You will then be able to enter a name for your enzyme, digest rules (see below) and select if you would like to find proteins that satisfy the rules at both ends. This option is grayed out for built-in enzymes.

Digest Rules: This is how you specify where your enzyme will cleave the protein between two amino acids to create peptides. The letter X denotes ‘any amino acid in this position’, while {set brackets} indicate any amino acid except the one in the brackets. Clicking on the “Advanced” button will open a new window which will allow you to be more specific with your digest rules.

7.1.3 General Options

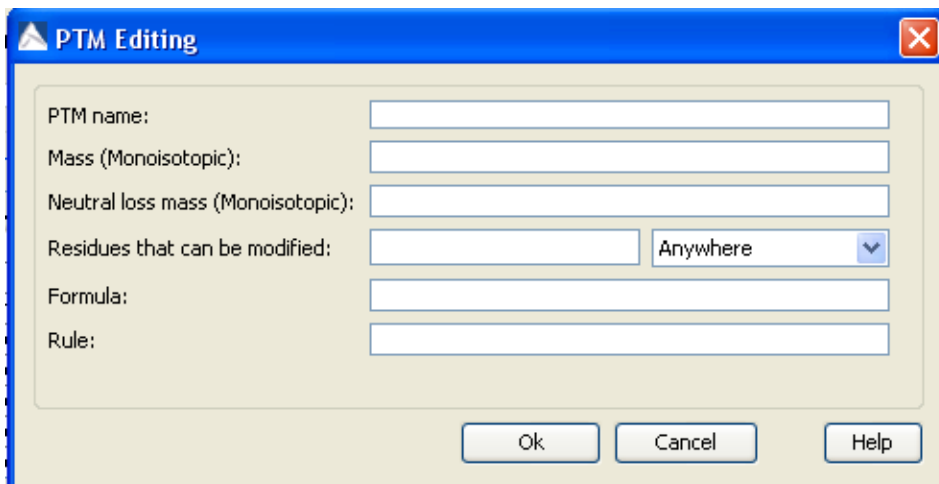
Max missed cleavages: determine the most missed cleavages to allow, internal to the peptide, in a *de novo* sequence. For instance, setting this to 2, and Trypsin as the enzyme, then PEAKS will return *de novo* sequences with up to 2 R's or K's internally.

Preprocess before auto de novo: PEAKS has its own built-in preprocessor for removing noise, centroiding and peak charge recognition from MS/MS data. Check this box to turn preprocessing on.

7.1.4 PTM Options

PTM options: This list tells PEAKS what kind of post-translational modifications to include in its analysis. Drag the desired PTM into either the “Fixed Modification” or “Variable Modification” box. If the desired PTM is not in the list, first check the “Show Unimod”

box to show additional PTMs. To create a new PTM click on the “New PTM” button. The following window will appear:



Fill in the following information:

Name: this name will appear in the PTM list for future use after it is saved.

Monoisotopic mass: the mass that the residue gains or loses as a result of the PTM. Enter this value numerically.

Neutral loss mass: the mass that the modified residue loses as a result of fragmentation. Ex. 28 would signify a loss of 28 Daltons. This is optional.

Chemical formula: the chemical formula of the PTM. This should correspond to the mass listed above. This is optional.

Residues that can be modified: Enter residues that can be modified anywhere, residues that can only be modified if they are at the N-or C-terminus or in the middle only.

Rule: Enter comments for reference. This is optional.

Please note that you can also configure your PTMs in the “Configuration” panel. See section [15.2.2 PTM Configuration](#) for more information.

Max variable PTM per peptide: To reduce uncertainty, limit PEAKS' *de novo* sequencing 'vocabulary', by restricting the number of variable PTM found on a peptide. Specify a number by typing it into the box. To lift such restrictions, type a very large number (longer than the length of the peptide).

7.1.5 Database Options

Database to search: Select from this drop-down list, one of the FASTA databases configured in PEAKS. To edit an already existing database, click on the “Edit Database” button. If the desired database is not in this list, click the "New Database" button. Note that you can also set up a new database in the “Database Configuration” window. The configuration window is the only place that that you can delete databases that you have created. For more information on setting up new databases see section [15.2.3 Database Configuration](#).

Taxonomy selection: This list displays the taxa you have chosen for your search. If the database selected has taxon information available, you can click on the “Set taxa” button. Otherwise, the whole database will be searched. The selections correspond to established hierarchy -- i.e. selecting 'Mammalia' will search all of 'horse, cow, rat, mouse, human, etc.

Paste FASTA sequences: If you already know the sequence of the protein(s) you are looking for, select “Paste fasta sequences” and paste the sequence in the space provided in fasta format. Alternatively, if you want search the same sequence regularly, it is recommended to simply create a small text file and configure it as a database for PEAKS.

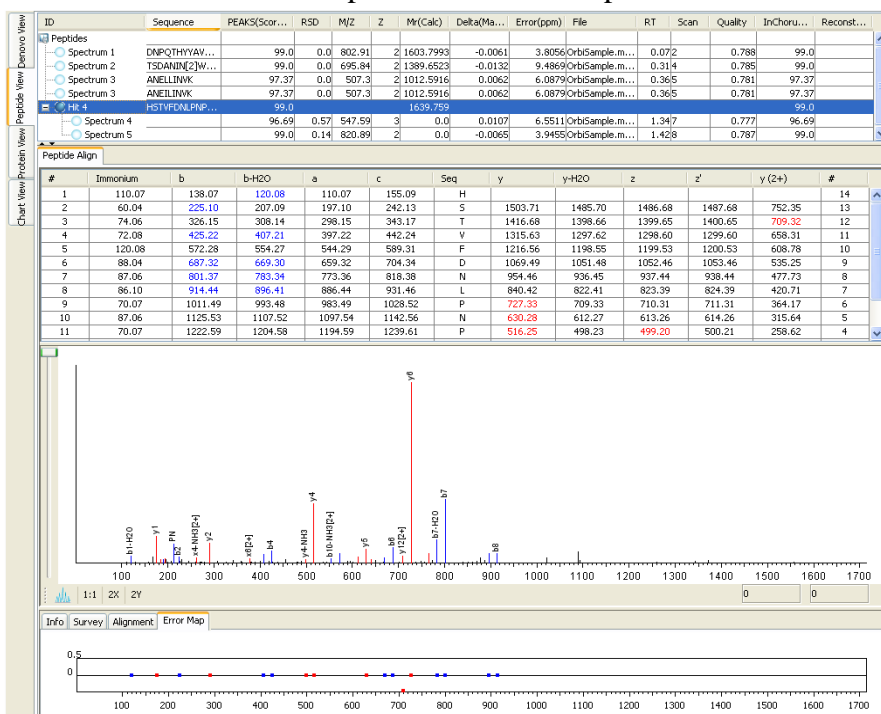
7.1.6 Advanced Options

PEAKS needs to have some *de novo* sequences before database searching since PEAKS uses sequence tags to perform database searching. As such the option of doing *de novo* prior to protein ID is presented here. In most cases, the same values for instrument, error, enzyme and PTM can be used in *de novo* and in protein ID, but you have the option of using one of your saved *de novo* parameter sets for the *de novo* portion. Select one from the drop-down list.

4) After setting up parameters, we can save them for future use. Click the "Save Parameters" button, and choose a name for future reference when prompted.

Any parameters that you save will be available in the drop-down list at the top of the window. To see what's inside, just select one, and the parameters boxes will be populated.

5) Press the “OK” button. If you have already performed *de novo* sequencing, the database search will commence automatically. If you have not previously performed *de novo* sequencing, the auto *de novo* process will appear first in the task queue. Once *de novo* sequencing is finished the database search will begin.



7.2 Protein Identification Results

7.2.1 Peptide View

Once PEAKS is finished searching the database, the “Peptide View” window will open by default:

The “Peptide View” window summarizes the results for each MS/MS spectrum. All peptides that match to each spectrum are displayed. By default the spectra are listed by ID in the “ID” column with the corresponding peptide sequence in the “Sequence column” beside. In certain cases, one peptide can correspond to more than one spectrum. These spectra are then listed in the “ID” column under a heading entitled “Hit”. Click on “+” to expand the view to see all of the spectra that can be matched by the same peptide.

The table below describes the contents of the columns in the “Peptide View Window”:

ID	A unique identifier for the MS/MS spectrum. <i>This differs from a scan number since we may have merged several scans together.</i>
Sequence	The amino acid sequence of the peptide. PTMs are listed in [square brackets].
Score	PEAKS’ probability score.
m/z	The measured mass/charge value, in Daltons, for the peptide.
z	The calculated charge value for the peptide
Mr (Calc)	The sum of the theoretical mass of the residues that form the identified peptide sequence from the database.
Delta (Mass)	The difference between Mr(Calc) and Mass, in Daltons.
Error (ppm)	The difference between Mr(Calc) and Mass, ppm.
File	The name of the file.
RT	Retention time (elution time) for the peptide as recorded in the scan header.
Scan	The scan number.
Quality	A value from 0 to 1 estimated from the spectrum to refer to spectrum quality. Attributes like signal to noise, total intensity, and spectrum tagging are used.
Scan Mode	Mode that the scan step was performed in.
Frag. Mode	Mode that the fragmentation step was performed in.

The columns themselves can be customized. Right click anywhere in the report and choose Toggle Column from the pop-up menu. The sub-menu that appears shows a checkmark in each of the columns that are currently showing. Click any one of them to show or hide a column. These settings will apply to all your reports.



Peptide Alignment

Click on the “Peptide Align” window. This will look very similar to the *de novo* results window. You will see the “Ion Table” which shows the proposed ions with their corresponding masses. To the right of the “Ion Table” is the “Error Plot” which displays the confidence that is assigned to each ion. The most confident results lie on the centerline. Clicking a cell or column in the Ion

Table highlights the corresponding points on the error plot and corresponding peaks on the spectrum.

Underneath the “Ion Table is the “Spectrum View Frame” which displays a graphical representation of the spectrum. The peptide that corresponds to the spectrum in the “Spectrum View Frame” is displayed in the “Input Sequence” box. Note that this is a drop-down menu so that you can select other peptides that have the same ID (if applicable). Scrolling over the spectrum will display a “tooltip” that will display the m/z ratio and the height/ intensity (as a percentage of 100) of that particular peak. Both the m/z ratio and the height of the peak can also found under the spectrum view on the right hand side.

To zoom either on the X or Y axes, select the “Zoom X” or “Zoom Y” buttons, respectively and then use the wheel on your mouse to move around the graph. Select the “Slide X” button and then use the wheel on your mouse to move around the graph. You must ensure that you are sufficiently zoomed in on the X axis to use the “Slide X” button. Selecting the “1:1” button will bring you back to the original image where you can see the entire spectrum on the screen.

You can use the profile  and peak  buttons to switch the spectrum view from profile mode to peak mode and vice versa. The scrollbar on the left acts to increase and decrease the intensity of the peaks, where the scrollbar on the right acts to zoom in to display the monoisotopic peaks.

Finally at the bottom of the screen is the “Spectrum Alignment Frame” which is used as a tool to navigate the “Spectrum View” frame. A blue bar along the horizontal axis of the alignment view indicates the range of the spectrum view in the “Spectrum View Frame”. This frame will show you how the proposed ions align with the spectrum. By default, the “Spectrum Alignment Frame” displays b-ion and y-ion. The b ions are shown right to left in blue, while the y ions are shown left to right in red. Clicking on the “Survey” tab will display the corresponding precursor ion spectrum. The buttons that appear in this section are the same as those that are explained above in the “Spectrum View Frame” section. Click on the “Error Map” tab. The m/z ratio is displayed on the y axis and the error is listed on the x axis in Daltons. The “Error Plot” displays the confidence that is assigned to each ion. The most confident results lie on the centerline. Clicking a cell or column in the Ion Table highlights the corresponding points on the error plot and corresponding peaks on the spectrum.

Peptide Align Peptide Details	
Select peptides for display:	sp Q29443 TRFE_BOVIN
Denovo DNPQTHYYAVVR Recon DNPQTHYYAVVR Homolog DNPQTHYYAVVR	
Matched peptides shown in blue, SPIDER matches shown in red	
1	MRFVAVRALLA CAVLGLCLAD PERTVWCTI STHEANKCAS FRENVLRIIE
51	SGPFVSCVKK TSHMDCIKAI SNNEADAVTL DGLVYEAGL KPNMLKPVVA
101	EFHGTV DNPQTHYYAVVR KDTDFKLNEL RGKKSCHTGL GRSAGWNIPM
151	AKLYKELPDP QESIQRAAAN FPSASCVPCA DQSSFFKLCQ LCAGKGTDKC
201	ACSNHEPYFG YSGAFKCLME GAGDVAFLVKH STVPDNLNFP EDRKNYELLG
251	GDNTRKSVDD TQRCYLAMPV SHAVVARTVG GKEDVIWELL NHAQEHFGKD
301	KPDNPFQLFQS PHGKDLLFKD SADGFLKIPS KMDFELYLGY EYVTALQMLR
351	ESKPPDSSKD ECMVKWCAIG HQERTKCDRW SGFSGGAIEC ETAENTECI
401	AKIMKGEADA MSLDGGYLYI AGKCGLVFVL AENYKTEGES CKNTPEKGYL
451	AVAVVKTSDA NINWNLLKDK KSCHTAVDRT AGWNIPMSLL YSKINCKFD

Peptide Details

Click on the “Peptide Details” tab. The following window will appear:

At the top of the “Peptide Details” frame is the accession number of the protein that corresponds to the peptide that you chose in the “Peptide View” window. If more than one protein matches a single peptide, you will be able to select these additional proteins using the drop-down menu.

Below this you will see a simple alignment between the original *de novo* sequence for this spectrum (if available), the peptide found in the database and the reconstructed sequence. Letters on a green background, and with vertical bars, indicate agreement. *Color* codes on the *de novo* sequence letters still indicate positional confidence.

Finally at the bottom of the window you will see the sequence of the selected protein and in blue you will see where the selected peptide matches the protein. The darker the blue, the more confident the match. The matched peptides will be shown in red if you have performed a SPIDER search which is discussed in the next section.

7.2.2 Protein View

Click on the “Protein View” tab. The following window will appear:

Protein View | Peptide View | De novo View

Accession	ID	Mass	Display	PEAKS(Score %)	Coverage(%)	Query matched	Marked	Description
DB Search	P00722 BGAL	1	116482.8		99.15	10.16	13	Beta-galactosida...
	Q29443 TRFE	2	77753.2		99.15	15.48	15	Serotransferrin p...
	P02769 ALBL	3	69293.55		99.1	14.99	14	Serum albumin pr...
	A7ZUE0 GLPK	4	56230.816		98.81	7.97	4	Glycerol kinase - ...
	P49064 ALBL	14	68659.56		98.52	7.57	5	Serum albumin pr...
	Q28522 ALBL	27	67881.09		91.41	5.67	3	Serum albumin pr...
	P00698 LYSC	29	16238.638		83.72	17.69	2	Lysozyme C prec...
	Q4UZR8 GLPI	31	55229.5		83.72	4.41	2	Glycerol kinase - ...
	P00330 ADH:	49	36823.152		83.69	5.17	2	Alcohol dehydrog...
	P07724 ALBL	52	68692.59		70.87	4.11	3	Serum albumin pr...
	Q9Y707 ACT:	60	41663.453		60.7	7.47	1	Actin-2 - Suillus b...
	Q1JP56 OPN:	63	54489.82		60.7	4.8	2	Melanopsin-like - ...
	P00706 LYSC	64	14507.404		60.7	11.63	1	Lysozyme C-3 - A...
	Q48K59 IAII	79	36881.0		60.7	3.85	2	1-aminocycloprop...
	Q6LNJ0 FTHE	81	61957.504		60.7	2.23	3	Formate-tetrahy...
	Q35ZR3 A1A	85	23182.424		60.7	5.45	1	Alpha-1-acid glyco...
	Q6DKE1 CYC	88	11695.467		60.7	9.52	1	Cytochrome c, te...
	P00017 CYC:	91	11682.467		60.7	9.52	1	Cytochrome c - A...
	Q03131 ERY:	124	365030.44		60.7	0.34	2	Erythronolide syn...
	Q92BH6 GLPK	126	55435.76		60.7	2.01	1	Glycerol kinase - ...

Modifications | Sequence Browser | Sequence Comparison

NCBI BLAST search of **P00722|BGAL ECOLI**

Link to retrieve entries containing this sequence from NCBI Entrez:

Accession/ID	Description
P00722 BGAL ECOLI	Beta-galactosidase - Escherichia coli (strain K12)

Peptides List:

ID	Sequence	PEAKS(...)	M/Z	Z	Mr(Calc)	Delta(...)	Error(p...)	File	RT	Scan	Quality	Scan M...	Frag. ...
Peptides													
	SpecQNNFNAYR	99.12	481.74	2	961.47296	0.0075	7.8082	TEST6_2...	19.523	1056	0.773	FT-ICR/O...	FT-ICR/O...
	Hit 1Y5QQ[1]...				1506.6885								
	SpecLTAAAC[1]...	99.12	477.72	2	953.4277	0.0022	2.3046	TEST6_2...	23.085	1189	0.755	FT-ICR/O...	FT-ICR/O...
	SpecAPLDNDI...	99.12	729.36	2	1456.7158	0.0104	7.1229	TEST6_2...	23.954	1224	0.786	FT-ICR/O...	FT-ICR/O...
	SpecHQQFF...	99.12	633.31	2	1264.6101	0.0046	3.6681	TEST6_2...	27.142	1356	0.784	FT-ICR/O...	FT-ICR/O...
	SpecGDFQFNISR	99.12	542.26	2	1082.5144	0.0089	8.232	TEST6_2...	27.56	1374	0.781	FT-ICR/O...	FT-ICR/O...
	Hit 6IGLNC[1]...				1471.7454								

Matched peptides shown in blue, red for SPIDER:

1 MTMITDSLAV VLQRRDWENP GVTQLNRLAA HPPFASWRNS EEARTDRPSQ

51 QLRSLNGEWR FAWFFAPEAV PESWLECDLP EADTTVVVPSN WQMHGYDAPI

The “Protein View” collects all the peptide identifications together, summarizes which proteins were present in the sample, and groups homologous proteins together. The same information is displayed in the Peptide View as in this Protein View; however, the results are organized to best enable us to evaluate at the protein level.

This view is helpful when building a summary that can be sent to a customer/collaborator. See chapter 13 for more details on exporting whole files or proteins of interest to an Excel file.

Index

The top section of this view (shown above) behaves like an index, listing each protein found in the sample. Very similar proteins, containing the same set or a subset of the matched peptides, are clustered together. To expand and collapse the full list of proteins within each cluster click the ‘+’ or ‘-’ sign respectively.

The table below describes the contents of the columns in the index:

Accession	The GI, accession or other unique identifier for this protein as recorded in the database that was
------------------	--

	searched.
Mass	The calculated mass of this protein
Display	A graphical coverage map. Blue areas represent parts of the sequence that have been explained by the identified peptides.
Score (%)	A value from 1 to 99 representing the confidence we have in this protein identification – calculated from the confidence on the ten best peptide hits for this protein, and normalized against the other identified proteins.
Coverage (%)	The number of amino acids in the protein sequence that have been explained by the identified peptides. Expressed as a percentage of the total length of the protein.
Query Matched	The number of spectra explained by matching to a peptide from this protein.
Marked	A multi-function checkbox. By default unchecked, but we can use this to select proteins for export or multiple sequence alignment.
Description	The part of the protein's header information as parsed from the database, usually it contains the name of the protein.

The columns themselves can be customized. Right click anywhere in the report and choose Toggle Column from the pop-up menu. The sub-menu that appears shows a checkmark in each of the columns that are currently showing. Click any of them to show or hide a column. These settings will apply to all your reports.

Sequence Browser

The “Sequence Browser” tab is selected by default. Clicking on a protein in the index will display the sequence of that protein below in the “Sequence Browser” panel. Clicking on the hyperlink of the accession number of the protein shown in blue will open a new window containing the webpage of the database that you searched for protein ID page in a new window.

NCBI BLAST search of [sp|Q29443|TRFE_BOVIN](#)

Link to retrieve entries containing this sequence from NCBI Entrez:

Accession/ID	Description
sp Q29443 TRFE_BOVIN	Serotransferrin OS=Bos taurus GN=TF PE=2 SV=1

There is also a “Peptides List” box which displays information about the peptides that matched to the selected protein. This list is identical to the “Peptide View” panel. Please see this section for more details.

Peptides List:

ID	Sequence	PEAKS(...)	RSD	M/Z	Z	Mr(Calc)	Delta(M...	Error(...)	File	RT	S...	Q...
Peptides												
● Spectrum 1	DNPQTHYYAVAVVK	99.0	0.0	802.91	2	1603.7993	-0.0061	3.8056 OrbiSam...		0.072		0.788
● Spectrum 2	TSDANIN[3]WNNLK	99.0	0.0	695.84	2	1389.6523	-0.0132	9.4869 OrbiSam...		0.314		0.785
■ Hit 4	HSTVFDNLPNPEDR	99.0				1639.759						
● Spectrum 5		99.0	0.14	820.89	2	0.0	-0.0065	3.9455 OrbiSam...		1.428		0.787
● Spectrum 4		96.64	0.57	547.59	3	0.0	0.0107	6.5511 OrbiSam...		1.347		0.777

Below the “Peptides List”, you will see the protein sequence with the matching peptide sequences in blue. The darker the blue, the more confident the match.

Matched peptides shown in **blue**, SPIDER matches shown in **red**

```

1  MRPAVRALLA CAVLGLCLAD PERTVRWCTI STHEANKCAS FRENVLRILE

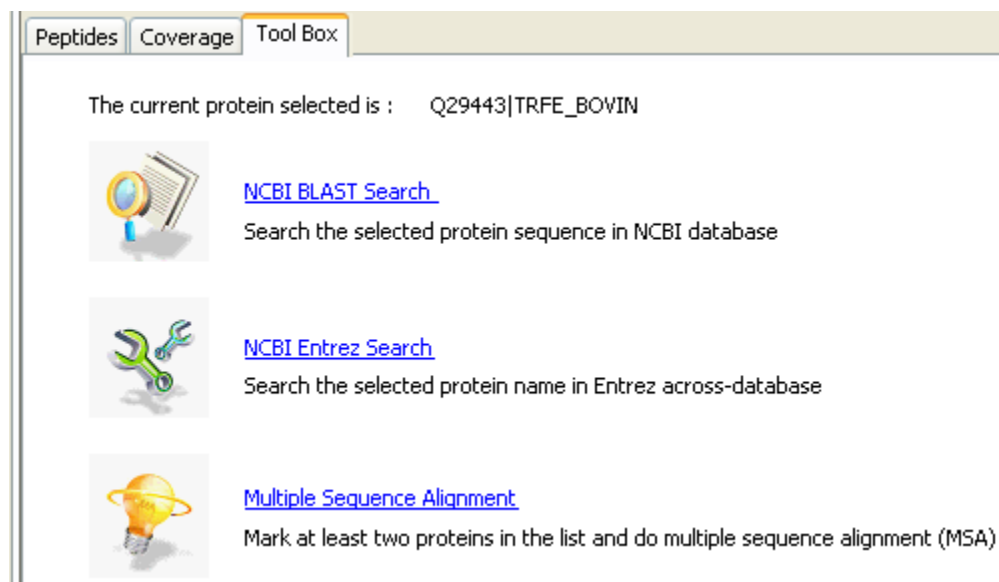
51  SGPFVSCVKK TSHMDCIKAI SNNEADAVTL DGGLVYEAGL KPNMLKPVVA

101 EFHGTDMPQ THYYAVAVVK KDTDFKLNEL RGKKSCHTGL GRSAGWNIPM

```

Tool Box

Click on the “Tool Box” tab:



The selected protein will be displayed at the top of this page. Click on NCBI Blast Search or NCBI Entrez Search to search for the selected protein sequence in the NCBI database or the Entrez database, respectively.

To perform a multiple alignment search mark two or more proteins in the above list of proteins, by clicking in their checkboxes (see example below).

PEAKS 1 [22-Jun-09 13:11] x

Accession	ID	Mass	Display	PEAKS(Score %)	Coverage(%)	Query matched	Description	Marked
DB Search								<input type="checkbox"/>
Q29443 TRFE_BOVIN	1	77753....		83.8	3.69	3	Serotransferrin - Bos t...	<input checked="" type="checkbox"/>
P00330 ADH1_YEAST	2	36691....		59.3	2.31	1	Alcohol dehydrogenas...	<input checked="" type="checkbox"/>

Then click on the “Multiple Sequence Alignment” link in the Toolbox window to generate the multiple sequence alignment in your web browser.

Identified peptides are highlighted in blue letters on the sequence. A more intense blue indicates a more confident match. The background *colors* indicate similarity between the sequences. A

dark background indicates regions where residues or nucleotides are identical in all sequences, a light background indicates similarity across some sequences, and lowercase letters on white background highlight differences. A dash - is displayed where a gap had to be introduced in one sequence to complete the alignment.

```

Q29443|TRFE_BOVIN      mrpavrallacavlgclcladpervrwctiStHEnKcasFrENvIRIlesgpfvscvkk
P00330|ADH1_YEAST      -----SipEtqKgviFyESHgKL-----

Q29443|TRFE_BOVIN      tsHmDcikaIsnnEADaVtLDgglVyeAGLkpnNLkpvvaeFHGtKdnPqthyyavavvk
P00330|ADH1_YEAST      -eHkDi--pVpkpKANEllIN---VkySGVchtDLha----WHGdwplP-----

Q29443|TRFE_BOVIN      kdtidfKLneLrGkKsch---tGLGrSa-GWnIpmaklykelpdpqesiqrAaanFfSaSCV
P00330|ADH1_YEAST      ----vKLplVgGhEgagvvvGMGeNvkGWkIg-----dyAgikWlNgSCM

Q29443|TRFE_BOVIN      pCadqssfpKlCQLcagkgtDKcaCsnHepyfGYS--GAF-KclmegAgdvAfVkhstvf
P00330|ADH1_YEAST      aC-----EyCEL-----gNEsnCp-HadlsGYThdGSFqQyatadAvqaAhIpqgTdl

Q29443|TRFE_BOVIN      dnLpnpedrknYeLLCgdnTrksvddyqecYlAMvpSHaVvartvvgkedviWellNhAq
P00330|ADH1_YEAST      aqVa-----pILCagiT-----vYkALksANlMagh-----WvaISgAa


```

8. SPIDER Search

After having obtained *de novo* sequences for peptides that are not in the database, it is important to look for a homologous peptide in the database (Han *et al.*, 2005). This will help you to learn more about the proteins in your sample. To search with SPIDER you must first have some good *de novo* sequences.

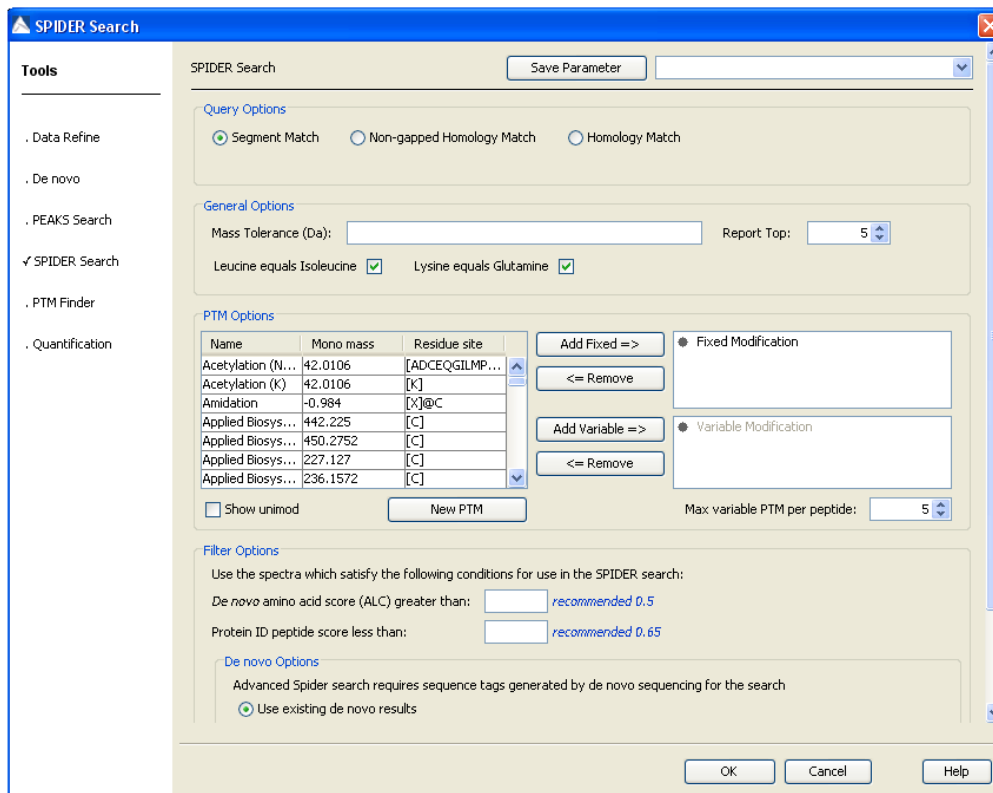
8.1 Setting up SPIDER Parameters

1) Select a data file or a Protein ID result from the “Project View” frame

2) Click the SPIDER icon on the toolbar 
Or

Choose SPIDER Search from the Tools menu.

When a Protein ID results file is selected, the “SPIDER Search Options” window will appear as seen below:



Tools

- . Data Refine
- . De novo
- . PEAKS Search
- ✓ SPIDER Search
- . PTM Finder
- . Quantification

SPIDER Search [Save Parameter]

Query Options

☒ Segment Match ☐ Non-gapped Homology Match ☐ Homology Match

General Options

Mass Tolerance (Da): [] Report Top: [5]

Leucine equals Isoleucine ☒ Lysine equals Glutamine ☒

PTM Options

Name	Mono mass	Residue site
Acetylation (N...)	42.0106	[ADCEQGILMP...]
Acetylation (K)	42.0106	[K]
Amidation	-0.984	[X]@C
Applied Biosys...	442.225	[C]
Applied Biosys...	450.2752	[C]
Applied Biosys...	227.127	[C]
Applied Biosys...	236.1572	[C]

☐ Show unimod [New PTM]

Fixed Modification

Variable Modification

Max variable PTM per peptide: [5]

Filter Options

Use the spectra which satisfy the following conditions for use in the SPIDER search:

De novo amino acid score (ALC) greater than: [] recommended 0.5

Protein ID peptide score less than: [] recommended 0.65

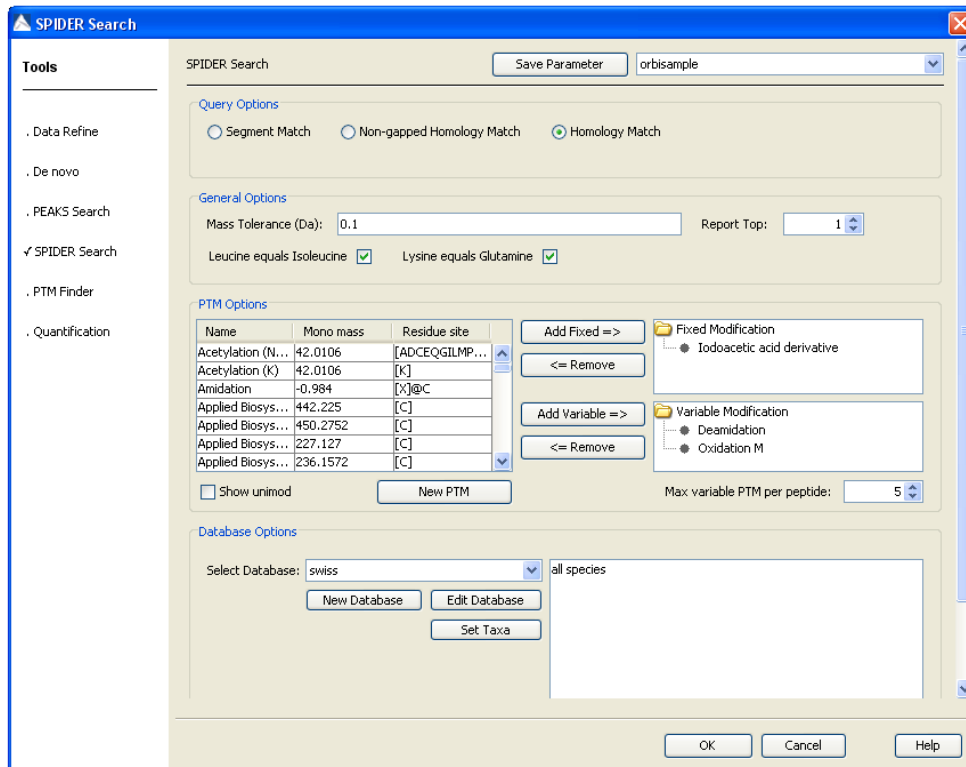
De novo Options

Advanced Spider search requires sequence tags generated by de novo sequencing for the search

☒ Use existing de novo results

[OK] [Cancel] [Help]

If you have selected a data file, the following window will appear:



Note that this window differs from the other window as it asks you to select a database. It also does not give you any filter options.

In this case, we assume you already have *de novo* results for the data file.

Selecting a database, provides sequence variations for SPIDER to search against *de novo* sequences.

3) The following section will describe the different options that you have when setting up the parameters for your SPIDER search.

8.1.1 Query Options

Choose a Query Type. They are, in order of increasingly rigorous analysis:

Segment Match: this is not a true mutation search; instead, it will insist that the mass of the peptide returned is the same as that of the *de novo* sequence.

Non-gapped Homology Match: this search will allow for transpositions and single point mutations but not insertions or deletions.

Homology Match: this is the most rigorous (but most resource intensive) search mode, taking into account all types of mutations and the positional confidence scores. A quick version of this is used to create the reconstructed peptides and to generate the final scores in each of the previous search modes. This is the only search mode that allows you to use variable modifications.

8.1.2 General Options

Amino acid selection: Choose if you would like PEAKS to consider Leucine equal to Isoleucine without a penalty in the score as well as whether Lysine should be equal to Glutamine without penalty.

Mass tolerance: Enter the amount of error (in Daltons) that PEAKS will allow for when determining the peptide sequences.

Number of peptides to report: Choose how many of the best homologous peptides should be displayed after searching

8.1.3 PTM Options

PTM Options List: The “PTM Options” list tells PEAKS what kind of post-translational modifications to include in its analysis. To view additional modifications, select the “Show unimod” box. If the desired PTM does not appear on the list or is different than what is listed, you can select the “New PTM” button and the “PTM Editing” window will open. Fill in the information pertaining to the PTM of interest. For a more in depth explanation of creating a new PTM, see section [7.1.4 PTM Options](#).

To select a PTM as Fixed or Variable, click within row of the desired PTM to highlight it, then click the button corresponding to its preferred designation, “Fixed Modification” or “Variable Modification”. If you incorrectly place a PTM in one of these sections, with it selected, simply select either Deselect, to have it removed from the list of selected modifications or Switch Type, to have it set as the other modification type. For example, using the Switch Type button, a PTM currently listed as fixed, will automatically be removed from the fixed modification list and enter the variable modification list, or vice versa.

Note that in previous versions of PEAKS, only fixed PTMs were allowed, however PEAKS version 5.1 allows variable PTMs as well when using the homology search.

Max variable PTMs: To reduce uncertainty, PEAKS' *de novo* sequencing 'vocabulary' can be limited by restricting the number of variable PTM found on a peptide. Specify a number by typing it into the box. To lift such restrictions, type a very large number (longer than the length of the peptide).

8.1.4 Filter Options

As the SPIDER search is computationally intensive, it is not recommended that you run all of your *de novo* sequencing peptides against the database; only those that cannot be well explained.

de novo score (A.A.) threshold: The SPIDER search requires a good sequence tag from *de novo* to be able to find good quality homologous proteins. Enter a value for the *de novo* score threshold. The recommended threshold is 0.5.

Peptide score threshold: Because there is no need to run SPIDER on peptides that already were found to have a good match during PEAKS protein ID, it is helpful to enter a peptide score threshold so that SPIDER will only be performed on peptides below the threshold. The recommended threshold is 0.65.

8.1.5 *de novo* Options

Because SPIDER requires a *de novo* sequence to find homologous proteins in the database, *de novo* sequencing will need to be performed first. If you have already done *de novo* sequencing, select the “I have already run *de novo*” button.

8.1.6 Database Options

Note that these options are only visible if you choose to run a SPIDER search on a data file rather than a PEAKS results file.

Database to search: Select from this drop-down list, one of the FASTA databases configured in PEAKS. To edit an already existing database, click on the “Edit Database” button. If the desired database is not in this list, click the “New Database” button. Note that you can also set up a new database in the “Database Configuration” window. The configuration window is the only place that that you can delete databases that you have created. For more information on setting up new databases see section [15.2.3 Database Configuration](#).

Taxonomy selection: This list displays the taxa you have chosen for your search. If the database selected has taxon information available, you can click on the “Set taxa” button. Otherwise, the whole database will be searched. The selections correspond to established hierarchy -- i.e. selecting 'Mammalia' will search all of 'horse, cow, rat, mouse, human, etc.

4) After setting up parameters, we can save them for future use. Click the “Save Parameters” button and choose a name for future reference when prompted. Any parameters that you save will be available in the drop-down list at the top of the window. To see what's inside, just select one, and the parameters boxes will be populated.

5) Press the “OK” button and the SPIDER search will begin.

8.2 SPIDER Results View

SPIDER will search the database for homologous peptides and attempt to consolidate these into protein hits as well. The result report will look much like the results for PEAKS Protein ID or inChorus searching.

Clicking on the “Peptide View” tab will display results that look very much like the results for PEAKS Protein ID. See section [7.2.1 Peptide View](#) for more details. Click on the “Peptide

Peptide AlignPeptide Details

Select peptides for display:sp|Q29443|TRFE_BOVIN

DenovoDNPQTHYTAVA/VK

ReconDNPQTHYAVAVVK

HomologDNPQTHYAVAVVK

Matched peptides shown in blue, SPIDER matches shown in red

1MRPAVRALLACAVLGLCLADPERTVRWCTISTHEANKCASFRENVLRLILE

51SGPPFVSCVKKTSHMDCIKAISNNEADAVTL DGLVYEAGL KPNMLKPVVA

101EFHGTDNPQTHYAVAVVKKDTDFKLNELRGKKSCHTGLGRSAGWNIFM

details” tab to see the SPIDER matches shown in red.

Letters on a green background with vertical bars, indicate agreement. Letters on a red background indicate sequencing error. Color codes on the *de novo* sequence letters still indicate positional confidence. Letters on a blue background indicate uncertainty or mutation. “+” signs represent more likely mutations. [Square brackets] indicate an equal

mass substitution, common non-critical *de novo* errors. <Angle brackets> indicate an equal mass substitution and a mutation.

When simply identifying exact peptides from the database using PEAKS Protein ID or inChorus, there is no need to reconstruct the ‘real’ sequence.

Clicking on “Protein View” will again yield a similar display as was seen for PEAKS Protein ID (see section [7.2.2 Protein View](#)). In the “Display” column, red region indicate areas of homology and potential mutation identified by SPIDER. Blue regions indicate areas of homology from the protein ID search, and purple indicates that the spectrum was identified by both SPIDER and protein ID.

SPIDER 11 [28-May-09 09:42] x							
Accession	ID	Mass	Display	PEAKS(Score %)	Coverage(%)	Query matched	Description
DB Search							
P00722 BGAL_ECOLI	1	116482.800		99.1	52.34	104	Beta-galactosidase OS=Escherichia coli (strain K1...
Q29443 TRFE_BOVIN	2	77753.200		99.1	66.19	207	Serotransferrin OS=Bos taurus GN=TF PE=2 SV=1
P02769 ALBU_BOVIN	3	69293.550		99.1	64.42	162	Serum albumin OS=Bos taurus GN=ALB PE=1 SV=
P00698 LYSC_CHICK	5	16238.638		99.1	57.14	33	Lysozyme C OS=Gallus gallus GN=LYZ PE=1 SV=1
P00330 ADH1_YEAST	10	36823.152		99.0	61.78	51	Alcohol dehydrogenase 1 OS=Saccharomyces cer.
Q9P4C2 ADH2_KLUMA	23	36968.246		98.7	32.76	21	Alcohol dehydrogenase 2 OS=Kluyveromyces mar.
A7ZUE0 GLPK_ECO24	27	56230.816		98.7	13.15	6	Glycerol kinase OS=Escherichia coli O139:H28 (str.
P62894 CYC_BOVIN	36	11703.532		98.7	50.48	24	Cytochrome c OS=Bos taurus GN=CYCS PE=1 SV=
P00704 LYSC_NUMME	41	14309.143		98.4	34.11	8	Lysozyme C OS=Numida meleagris GN=LYZ PE=1
P00761 TRYP_PIG	74	24409.480		95.9	26.41	6	Trypsin OS=Sus scrofa PE=1 SV=1
A9B381 GLPK_PETMO	76	55398.965		90.4	12.45	4	Glycerol kinase OS=Petrotoga mobilis (strain DSM
Q6DK11 CYC2_XENLA	89	11695.467		81.8	20.95	8	Cytochrome c, testis-specific OS=Xenopus laevis
Q35ZR3 A1AG_BOVIN	94	23182.424		80.5	12.38	4	Alpha-1-acid glycoprotein OS=Bos taurus GN=OR.

Click on the “Coverage” tab. The highlighted *colors* are the same as those shown in the “Display” column above.

Peptides		Coverage	Tool Box
DB peptides shown in blue SPIDER matches shown in red Overlapping matches in magenta			
1	MTMITDSLAV VLQRRDWENP GVTQLNRLAA HPPFASWRNS	BEARTDRPSQ	
51	QLRSLNGEWR FAWPPAPEAV PESWLECDLP EADTVVVPSP	WQMHGYPAPI	
101	YTNVTYPIIV NPPFVPTENP TGCYSLTFNV DESWLQEGQT	RIIFDGVNSA	
	1		
151	PHLWCNGRWV GYGQDSRLPS EFDLSAPLRA GENRLAVMVL	RMSDGSYLED	
	1		
201	QDMWRMSGIF RDVSLHKPT TQISDFHVAT RFNDDFSRAV	LEAEVQMCGE	
251	LRDYLRLVTVS LWQGETQVAS GTAPFGGEII DERGGYADRV	TLRLNVENPK	
	1		
301	LWSAEIPNLY RAVVELHTAD GTLIEAEACD VGFRVRIEM	GLLLNGKPL	
	2		
351	LIRGVNRHEH HPLHGQVME QTMVQDILLM KQNNFNAVRC	SHYPNHLWY	

9. PTM Finder

9.1 Setting up PTM Finder Parameters

1) Select a Protein ID results file to perform a PTM finder search on. Note that you cannot perform protein ID on a raw file or *de novo* results.

2) Click the PTM icon on the toolbar 

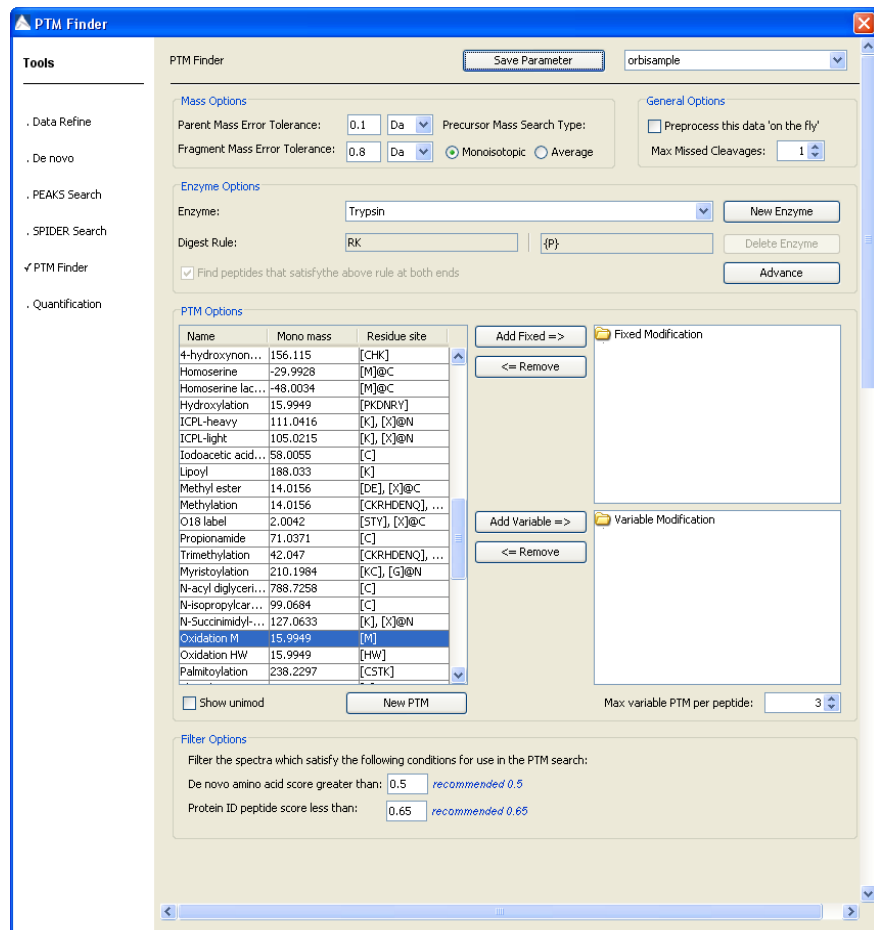
Or

Select “PTM Finder” from the Tools menu.

The “PTM Finder Options” window will appear:

The parameters are the same as you used when performing protein ID (7.1 Setting up Protein Identification Parameters) with the exception of the filter options found at the bottom of the window. As PTM Finder searches tend to be computationally intensive, PEAKS will only look at *de novo* sequencing results that are above the amino acid score threshold and below the peptide score threshold that you input.

de novo score (A.A.) threshold: The PTM finder requires a good sequence tag from *de novo* to be able to find good quality homologous proteins.



PTM Finder

Save Parameter orbisample

Mass Options

Parent Mass Error Tolerance: 0.1 Da Precursor Mass Search Type: ☐ Monoisotopic ☐ Average

Fragment Mass Error Tolerance: 0.8 Da

Enzyme Options

Enzyme: Trypsin New Enzyme

Digest Rule: RK (P) Delete Enzyme

☒ Find peptides that satisfy the above rule at both ends Advance

PTM Options

Name	Mono mass	Residue site
4-hydroxynon...	156.115	[CHK]
Homoserine	+29.9928	[M]@C
Homoserine lac...	+48.0034	[M]@C
Hydroxylation	15.9949	[PKDNRY]
ICPL-heavy	111.0416	[K], [X]@N
ICPL-light	105.0215	[K], [X]@N
Iodoacetic acid...	58.0055	[C]
Lipoyl	188.033	[K]
Methyl ester	14.0156	[DE], [X]@C
Methylation	14.0156	[CKRHENQ], ...
O18 label	2.0042	[STY], [X]@C
Propionamide	71.0371	[C]
Trimethylation	42.047	[CKRHENQ], ...
Myristoylation	210.1984	[KC], [G]@N
N-acyl diglyceri...	788.7258	[C]
N-isopropylcar...	99.0684	[C]
N-Succinimidy...	127.0633	[K], [X]@N
Oxidation M	15.9949	[M]
Oxidation HW	15.9949	[HW]
Palmitoylation	238.2297	[CSTK]

☐ Show unimod New PTM

Fixed Modification

Variable Modification

Max variable PTM per peptide: 3

Filter Options

Filter the spectra which satisfy the following conditions for use in the PTM search:

De novo amino acid score greater than: 0.5 recommended 0.5

Protein ID peptide score less than: 0.65 recommended 0.65

Enter a value for the *de novo* score threshold. The recommended threshold is 0.5.

Peptide score threshold: Because there is no need to run the PTM finder on peptides that were already found to have a good match during PEAKS protein ID, it is helpful to enter a peptide score threshold so that SPIDER will only be performed on peptides below the threshold. The recommended threshold is 0.65.

9.2 PTM Finder Results View

The results from a PTM finder search are identical to those seen in a PEAKS Protein ID search. Please see section [7.2 Protein Identification Results](#) for more information on the PEAKS Protein ID search results.

Chapter 10

10. inChorus Meta Search

inChorus protein identification can call upon several search engines (PEAKS protein ID, X!Tandem, OMSSA, Mascot and Sequest) for protein identification and will then compare and summarize the results from the different search engines in one single report.

Please note that if you wish to launch either Mascot or Sequest for an inChorus search, these engines are not internally provided with PEAKS and may need to be purchased by the respective vendors; however, licenses of these engines are not required if the result files have already been obtained. To set up your search engine preferences, see section [15.1.3 Search Engine Preferences](#).

10.1 Setting up inChorus Parameters

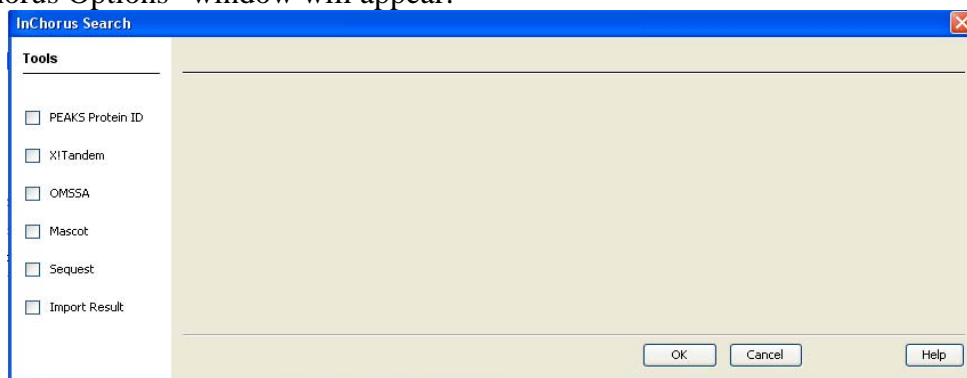
1) Select the orbisample.mzxml file

2) Click the “inChorus Search” icon on the toolbar 

Or

Select “inChorus Search” from the Tools menu.

The “inChorus Options” window will appear:



3) First select each of the protein identification tools that you would like to use by putting a checkmark in their respective checkboxes. Search parameters for each program can be set by selecting the name of the search engine.

The option screens for each of the search engines available to inChorus are designed to work in the same way as options screens from the original programs. For help in setting search

parameters for each program, please refer to that program's user manual. For help with PEAKS Protein ID, please see section [7.1 Setting up Protein Identification Parameters](#).

10.1.1 Importing Existing Results

PEAKS inChorus reads X!Tandem .xml files, OMSSA .omx files, Mascot .dat files and Sequest .srf files. When importing third party results files, please make sure that the scan number model in the results file is consistent with the one in PEAKS. PEAKS uses original data information to compute the inChorus score. When you run an X!Tandem search with command line, you need to turn on the option of “-w” in order to export data information into X!Tandem .xml files.

To import existing results, check the “Import Result” checkbox and select “Import Result”. The following window will open.

The screenshot shows the 'InChorus Search' window with the 'Import Result' tab selected. On the left, under 'Tools', the 'Import Result' checkbox is checked. The main area is divided into two sections: 'Database' and 'Import Search Result'. The 'Database' section has a dropdown for 'Database to search' set to 'swiss' and a text box for 'Taxonomy selections' set to 'all species', with a 'Set Taxa' button below. The 'Import Search Result' section contains four rows, each with a text input field and a 'Browse' button: 'Import X!Tandem .xml file', 'Import Omssa .omx file', 'Import Mascot .dat file', and 'Import Sequest .srf file'. At the bottom right are 'OK', 'Cancel', and 'Help' buttons.





Although it is not necessary for the various search engines to use the same database in an inChorus search, it is necessary to have a unified database for an inChorus search that includes imported results. Select the database that you would like to use from the drop-down list. The inChorus search will be performed on all species in the database unless specified by the user. If this database does not appear in this list, refer to 15.2.3 Database Configuration, to configure your databases. To specify which taxa you would like to search, click on the “Set Taxa” button.

To import your file, click the “Browse” button that is found beside the appropriate search engine. Find the file that you would like to import and click “Open”. Once you have selected the file(s)

that you would like to import and have selected the options for any other search engine searches you would like to perform, click “OK”.

10.2 inChorus Results View

When the inChorus search is complete the “Project View” panel should contain a separate results file for each search engine that you selected as well as an inChorus report that combines the results from the multiple search engines. See an example below:

...  PEAKS 3 [12-Jan-09 14:25]
...  XTANDEM 4 [12-Jan-09 14:26]
...  OMSSA 5 [12-Jan-09 14:28]
...  INCHORUS 6 [12-Jan-09 14:28]

10.2.1 *de novo*, Peptide and Protein Views

Each results file for the 3rd party search engines looks very similar to the PEAKS protein ID results file (7.2 Protein Identification Results) with a few small differences. Firstly, there is no “*de novo* View” or “Chart View”, and secondly, the scoring will be specific to that search engine. For example, the score for OMSSA is listed as an E-value. For more information about different scoring methods, refer to the user manual of the 3rd party search engines.

The inChorus search report also looks very similar to the 3rd party search engine results; however there are no “*de novo* View” results available. The *de novo* sequencing results that are found in the “*de novo* View” are only those that correspond to results that have been identified by one of the search engines in the inChorus search.


Chapter 11

11. Creating a High-Throughput Work Flow

In some situations, you may have many data sets that you wish to process all at once, with the same parameters. PEAKS 5.1 allows you to do this type of work with minimal effort on your part. By setting up a work flow, you can start a batch process of several data files and not worry about it until it is finished.

It is important to note that all the files you load will be processed in exactly the same way, using exactly the same parameters. If you want to do some differently than others, we must set up another work flow.

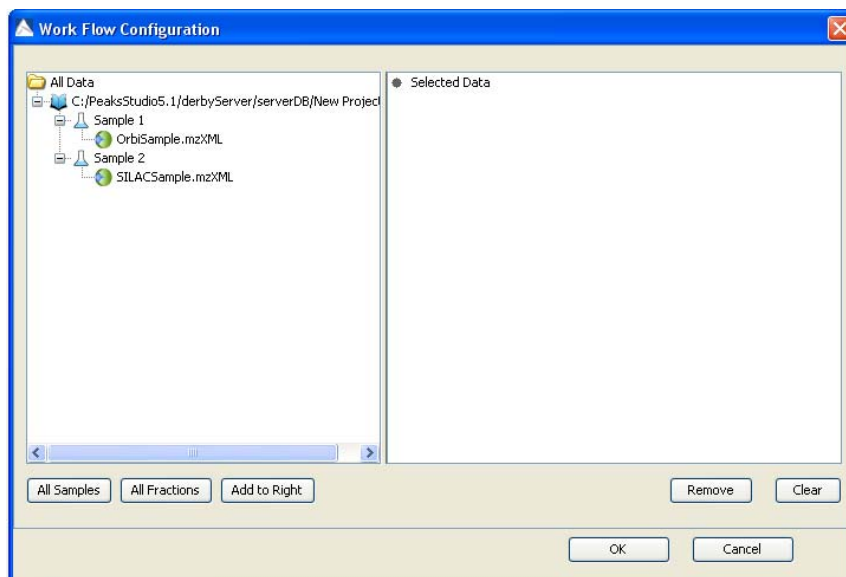
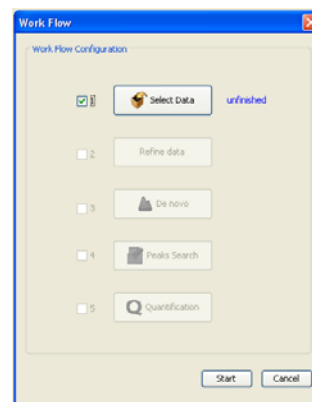
11.1 Setting up the Work Flow Parameters

Click the “Work Flow” icon on the toolbar 

The “Work Flow Configuration” window will appear:

The steps will appear as “unfinished” until they have been completed, at which time they will be labeled “finished”.

Click the “Select Data” button. The following window will open:



Only projects that are open in the “Project View panel” can be selected for analysis with the work flow function. To select which files/samples you would like to analyze, select the individual file/sample or click the “All Samples” or “All Fractions” button and then click the “Add to Right” button to transfer the samples/files to the “Selected Data” panel on the right hand side. Use the “Remove” and “Clear” buttons to remove selected files/samples or all files/samples, respectively from the “Selected Data” panel. Click “OK” to move to the next step.

Check the boxes beside the functions that you would like to perform with the work flow. Note that PEAKS requires *de novo* sequencing results to perform Protein ID. The parameter windows that open when you select the various functions in the work flow are identical to those that appear when you perform each function individually. Please refer to the chapters on each individual function if you require more details on setting up the parameters.

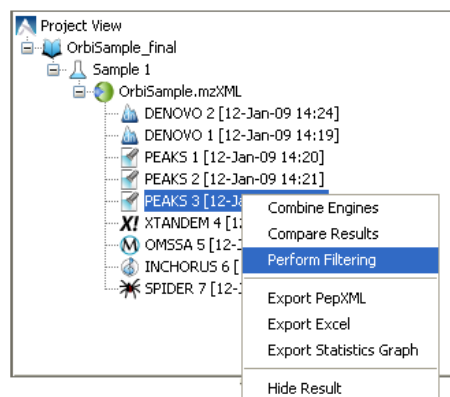
Chapter 12

12. Filtering Your Results

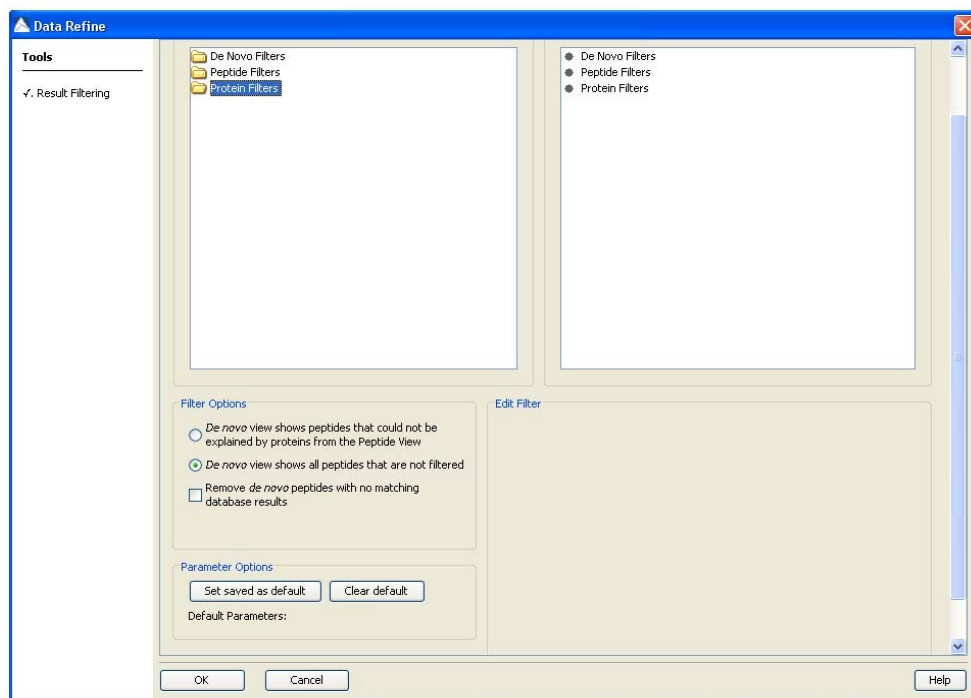
PEAKS 5.1 provides you with an exhaustive list of all proteins and peptides that can be found in a sample. However, as many researchers have their own criteria of what information is required in their report and what is an acceptable result, PEAKS 5.1 provides the necessary filtering tools that enable filtering out the less critical information, leaving just the essentials.

12.1 Setting Filter Parameters

Click on the time and date stamp associated with the result that you would like the filter. Once the report loads, click the right button on your mouse and select “Perform Filtering”.

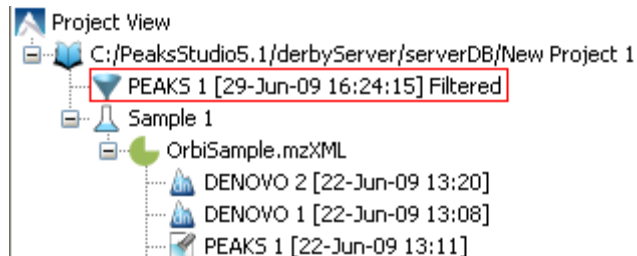


The following window should appear:



Note: PEAKS will automatically create a new filtered results node in the project view that will save your filter choice every time you click “OK”.

The filter node relies upon the original result so do not delete the original result if you want the filtered result to remain.

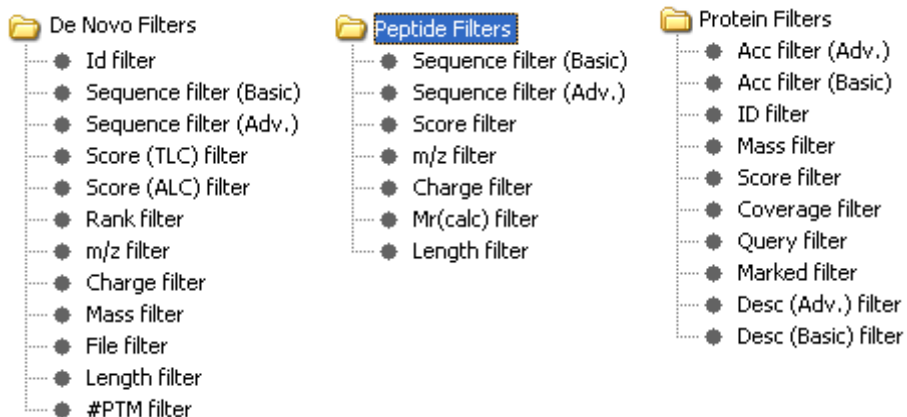


12.1.1 Possible Filters/ Selected Filters/ Edit Filter

The filters are grouped into three basic types to reflect what they act on:

- *de novo* filters act to remove proposed *de novo* sequences
- Peptide Filters act to remove peptides found in the database from the report
- Protein filters act to remove proteins from the report

To see the available filters for each level of filtering, double click on the appropriate folder in the “Possible Filters” frame. See the options in each folder below:



Choose a filter from the ‘Possible Filters’ list on the left by clicking on it. Options for this filter will appear in the “Edit Filter” frame. Once you have set the options that you would like for the filter in the “Edit Filter” frame, drag the filter that you would like into the “Selected Filter” list on the right hand side. Click “OK” to apply the filter that you have selected to the current file.

If you would like to add another filter, you can repeat the process, continuing to add as many filters as necessary. In this way it is also possible to have two filters on the same property; we can set a range of protein mass, for instance, by applying one filter on the upper bound of the mass and adding another filter to be the lower bound of the mass. We can also have more complex filters that involve multiple properties.

For example, let's say that you want to show only proteins with more than one high scoring (greater than 60% score) peptide, a standard requirement for publication. Double click on the "Peptide Filters" folder. Select "Score Filter". Edit the filter to select peptides that have a score that is greater than 60% in the "Edit filter" frame.

Score (%):
Filter peptides based on their Score(%)

60

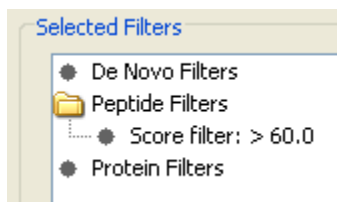
Options:

☐ Equal to

☒ Greater than

☐ Lesser than

☐ Not equal to



Now drag "Score filter" from the "Possible Filters" frame to the "Selected Filters" frame.

12.1.2 Filter Options

The filters cascade through each view in the multi-part report. For instance, removing a peptide from the database search results – the Peptide View list – will cause the *de novo* sequence for that peptide to be removed from the "de novo View" tab but will not affect the "Protein View" list. Filtering out a protein (by mass, for example) will remove it from the "Protein View" list and will remove all peptides associated with that protein from the "Peptide View" list as well as from the "de novo View" list. The manner in which the "de novo View" is linked can be specified by the user using the options in the "Filter Options" frame. See the figure below for more information.

Option Selected (on filter pane tab)

de novo View (represented in blue)

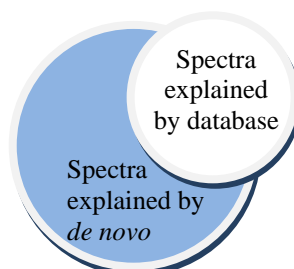
Linking Filters

Filter Options

☒ De novo view shows peptides that could not be explained by proteins from the Peptide View

☐ De novo view shows all peptides that are not filtered

☐ Remove de novo peptides with no matching database results



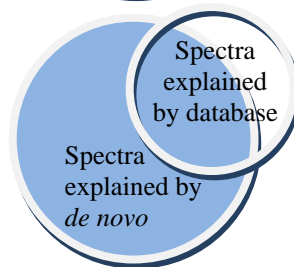
Filtering out (removing) a *de novo* result does not affect what is displayed in the Peptide View tab or Protein View tab.

Filter Options

☐ De novo view shows peptides that could not be explained by proteins from the Peptide View

☒ De novo view shows all peptides that are not filtered

☐ Remove de novo peptides with no matching database results



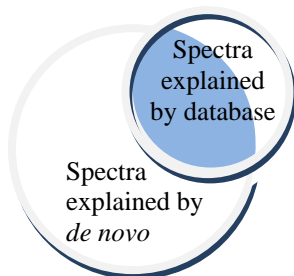
Filtering out (removing) a *de novo* result also removes that spectrum from consideration in the Peptide View tab and Protein View tab.

Filter Options

☐ De novo view shows peptides that could not be explained by proteins from the Peptide View

☒ De novo view shows all peptides that are not filtered

☒ Remove de novo peptides with no matching database results

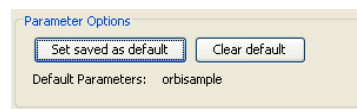


Again, filtering out (removing) a *de novo* result also removes that spectrum from consideration in the Peptide View tab and Protein View tab.

12.1.3 Parameter Options

Filter sets can be saved and re-used between sessions by clicking the “Save Parameter” button that is found at the top right hand corner of the “Filter Parameters” window.

You may prepare your results the same way each time; in which case it makes sense to set up a filter that will be automatically applied each time we load a report. Select a filter from the list of saved filters from the drop-down menu found at the top right hand corner of the “Filter Parameters” window. Click on the “Set saved as default” button. This filter will be displayed in the “Parameter Options” frame (as seen below) and will be applied automatically just after a report is loaded. Be careful, if your default filter is very stringent, it can sometimes remove everything! To remove a default filter, press the Clear Default Button at the bottom of the Filter Pane.



Each filter can be applied several times over. So it can get a little complex. To illustrate, here are a few examples:

- 1) Goal: Show proteins that have two high-scoring hits:
 - Add the Protein Filter called “Query” and in the “Edit Filter” section choose ‘greater than’ and type ‘1’ in the box (without the quotes). This will remove any ‘one hit wonders’.
 - Add the Peptide Filter called ‘Score’ and in the “Edit Filter” section choose ‘greater than’ and type ‘50’ in the box (without the quotes).
- 2) Goal: Find a protein that contains the word ‘human’ or ‘rat’ in the database entry’s description, but not Keratin or Trypsin.

- Add the Protein Filters called ‘Desc’
- In the “Edit filter” section, you are required to type in a regular expression (regex). This allows you to use wildcards.

Wildcard	Meaning	Example
.*	<i>“Anything of any length”</i>	.*human.* Will find anything that contains the word ‘human’, with anything before and anything after.
	<i>“Or” (use brackets)</i>	.*(human rat RAT).* Will find anything that contains the word ‘human’, or the word ‘rat’ or the word ‘RAT’, with anything before and anything after.
?!	<i>“Not” (use brackets)</i>	(?!.*(Keratin Trypsin).*).*(human rat).* will find anything containing human or rat but not Keratin or Trypsin
[]	<i>“Any of these characters”</i>	.*([Hh]uman [Rr]at).* will find anything containing the words Human, human, Rat or rat.

- Type in the regex: `(?!.*(Keratin|Trypsin).*).*(human|rat).*` and press the Enter key. If PEAKS confirms that this is a valid regular expression, it will put a check in the 'Valid Java Regex' box .

3) Goal: Setting a protein mass range

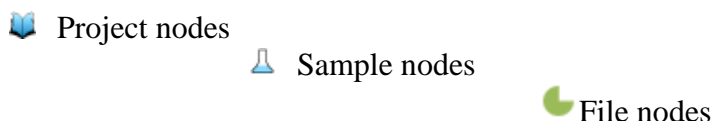
If we know the approximate mass of the proteins you are interested in, you can eliminate all proteins that are not close in mass.

- Add two filters: "Protein Filters: Mass >12000" *and* "Protein Filters: Mass < 32000".

13. Complex Analysis

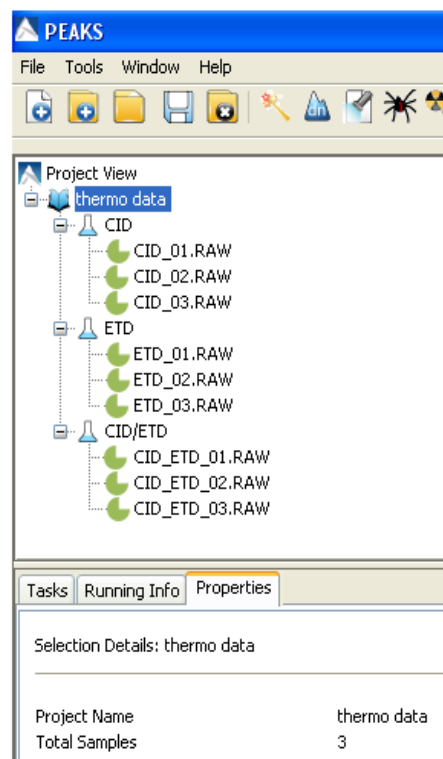
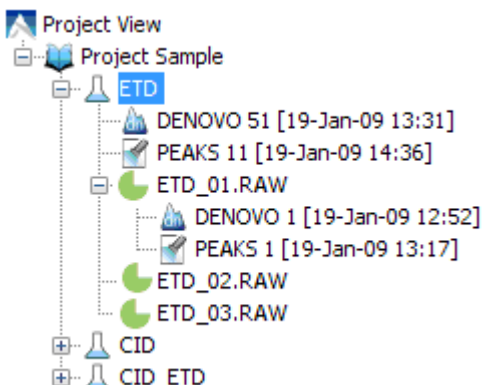
13.1 Creating a project for complex system

PEAKS 5.1 is able to analysis MS data from very complex systems. The data analysis scheme is organized as follows:



Below is an example of a project that contains three samples that were generated using the following fragmentation methods: ETD, CID, and CID/ETD. Each sample has three files.

13.2 Integrating data analysis



Within a project, the data is analyzed either file by file or sample by sample. By selecting a sample, the operation applies to all files in the sample. It means all spectra from different files are processed in a single run. The result node is displayed at same level as selected data node.

PEAKS 1 is the PEAKS Protein ID result for file ETD_01.RAW, whereas PEAKS 11 is the PEAKS Protein ID result for all three fractions of the ETD sample. Note that the result of a sample may not be the sum of the results of all files in the sample.

14. Exporting Data/Reports and Printing

PEAKS 5.1 allows you to create reports to share with collaborators, colleagues and clients. The reports are available in HTML or Microsoft Excel (.xls) formats and follows a 'What You See Is What You Get' philosophy. All the information you see on screen in PEAKS 5.1 will appear in the exported report. For this reason, it is important that we complete results filtering and toggling columns before exporting a report.

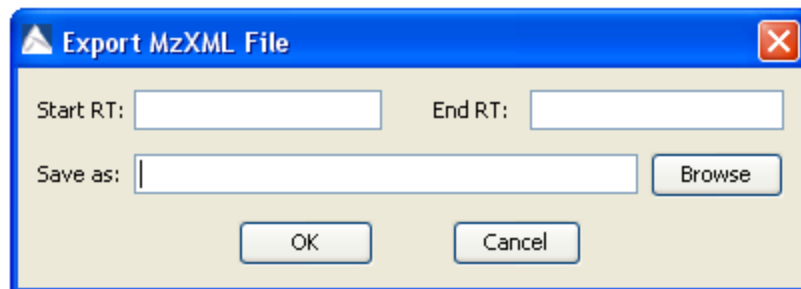
14.1 Export Data in .mzxml or .mgf

In order to export your data file in .mzxml or .mgf right click on the data file that you wish to export.

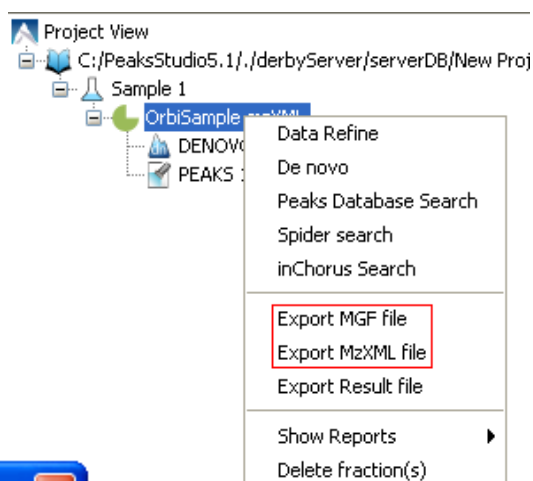
Click "Export MGF File". A window will open that will prompt you to enter a name and a location for the file. Click "Export".

or

Click "Export MzXML File". The following window will open:

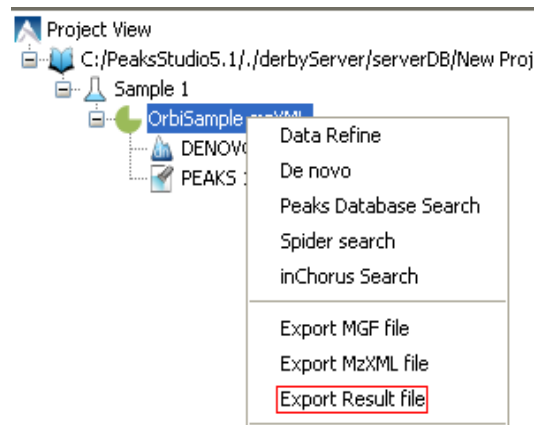


Enter the starting and ending retention times in the appropriate boxes. Then click the "Browse" button to select a destination to save your file.



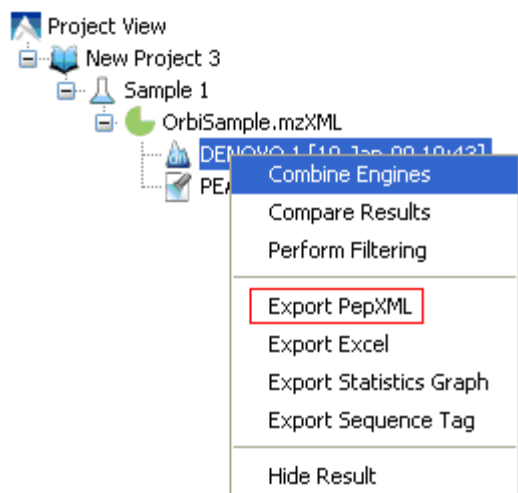
14.2 Export Results File

To export the results for a selected data file, right click on the data file and select "Export Result file". All of the PEAKS results file generated for the selected data file will be displayed in Excel format.

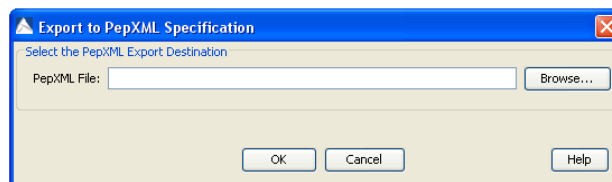


14.3 Export Peptide Results in PepXML Format

In order to export your PEAKS results file in PepXML right click on the results file that you wish to export and select “Export PepXML”.



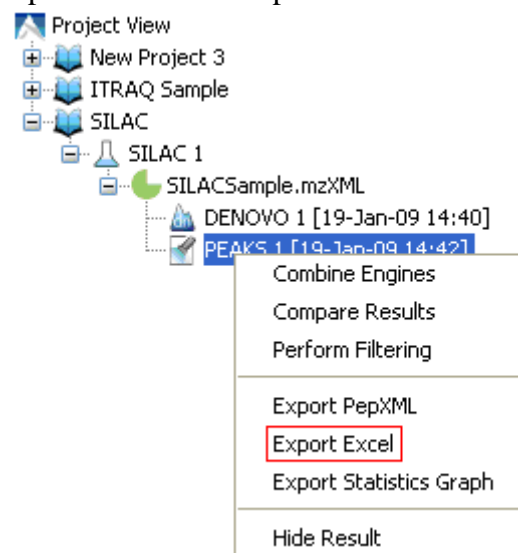
The following window will appear:



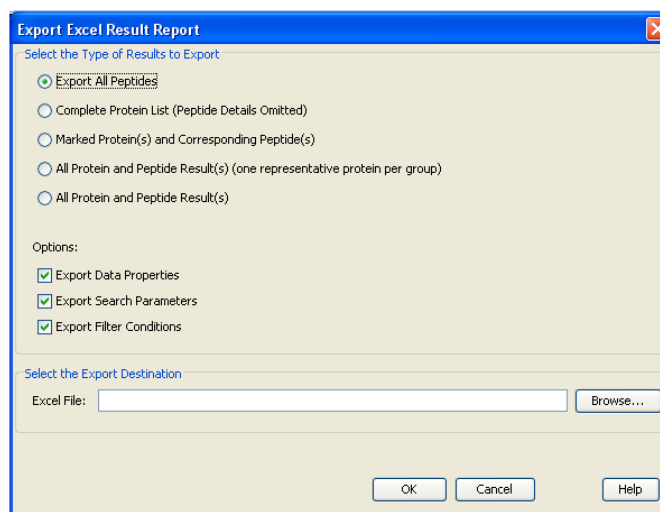
Browse through your computer to select the location that you would like to place the exported PepXML file. Then click “OK”.

14.4 Export Results in Excel Format

In order to export your PEAKS results file to Excel right click on the results file that you wish to export and select “Export Excel”.

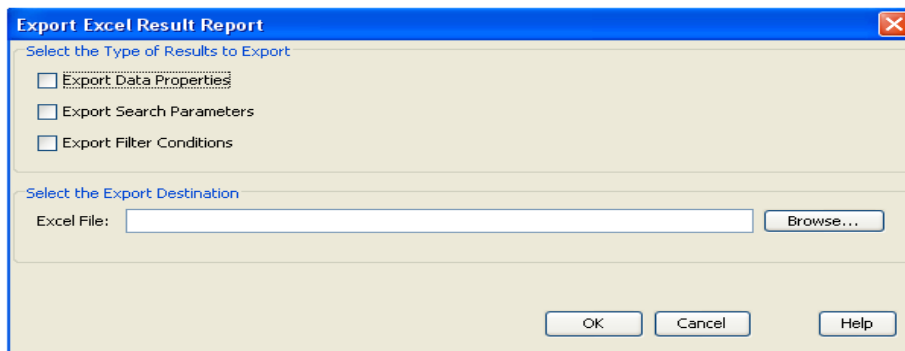


The following window will appear:



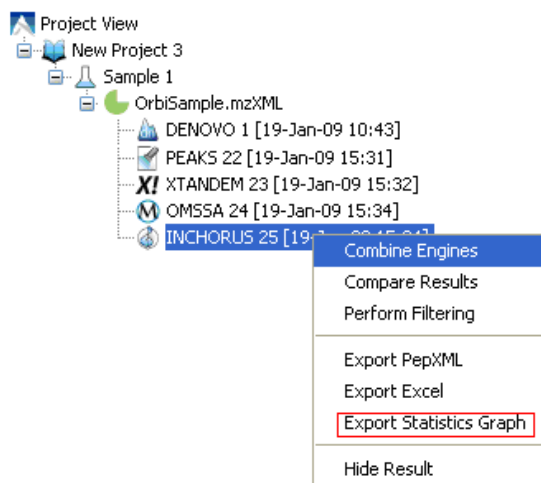
If you would like to export all of the protein and peptide results select the “All Protein and Peptide Result(s)”, otherwise select one of the other options where you can limit which results are exported. Select the appropriate boxes if you would like to export “Data Properties”, “Search Parameters” and “Filter Conditions” to Excel. Finally you need to select the “Export Destination” by clicking the “Browse” button. Then click “OK”.

If you export *de novo* results to Excel, the *de novo* sequencing results will be exported and you have the choice to also export “Data Properties”, “Search Parameters” and “Filter Conditions”. See below:



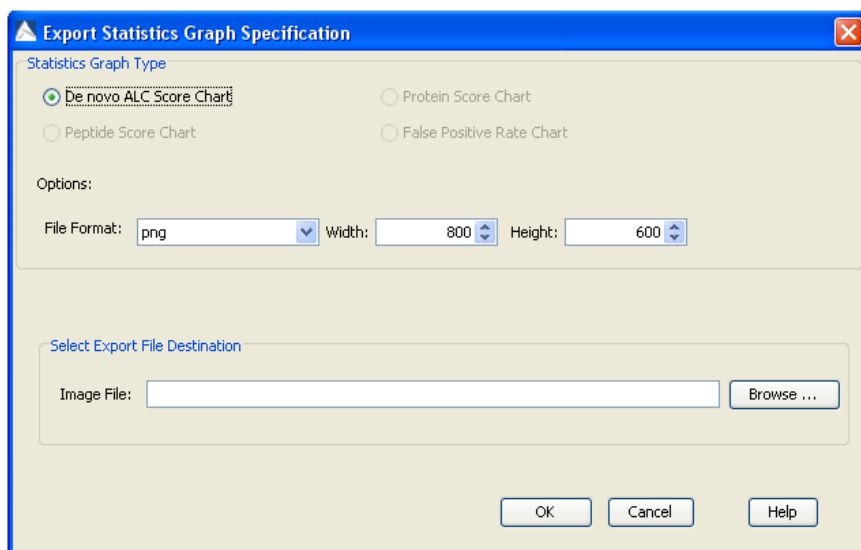
14.5 Print Tables and Graphs for Publication

In order to export an image file, right click on the results file that contains the appropriate image file and select “Export Statistics Graph”.

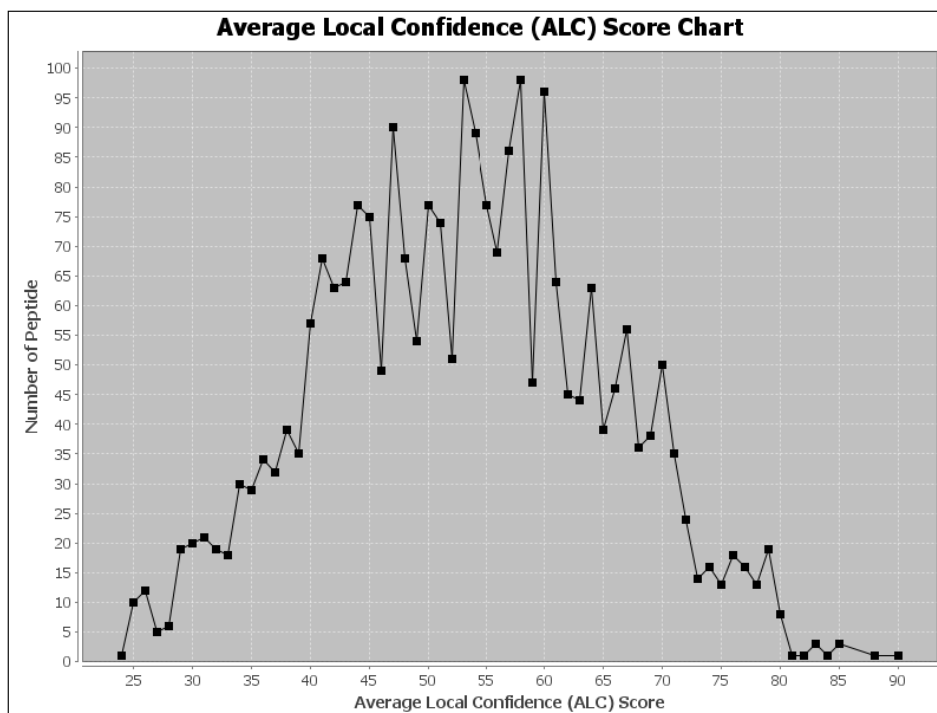


14.5.1 *de novo* Image Files

Select a *de novo* results file. Right click and select “Export Statistics Graph”. The following window will appear:



The “*de novo* ALC. Score Chart” option will be selected in the “Statistics Graph Type” panel. Select a file format and height/width for your chart and browse your computer to select a destination. Beside is an example of the output:



14.5.2 Protein ID Image Files

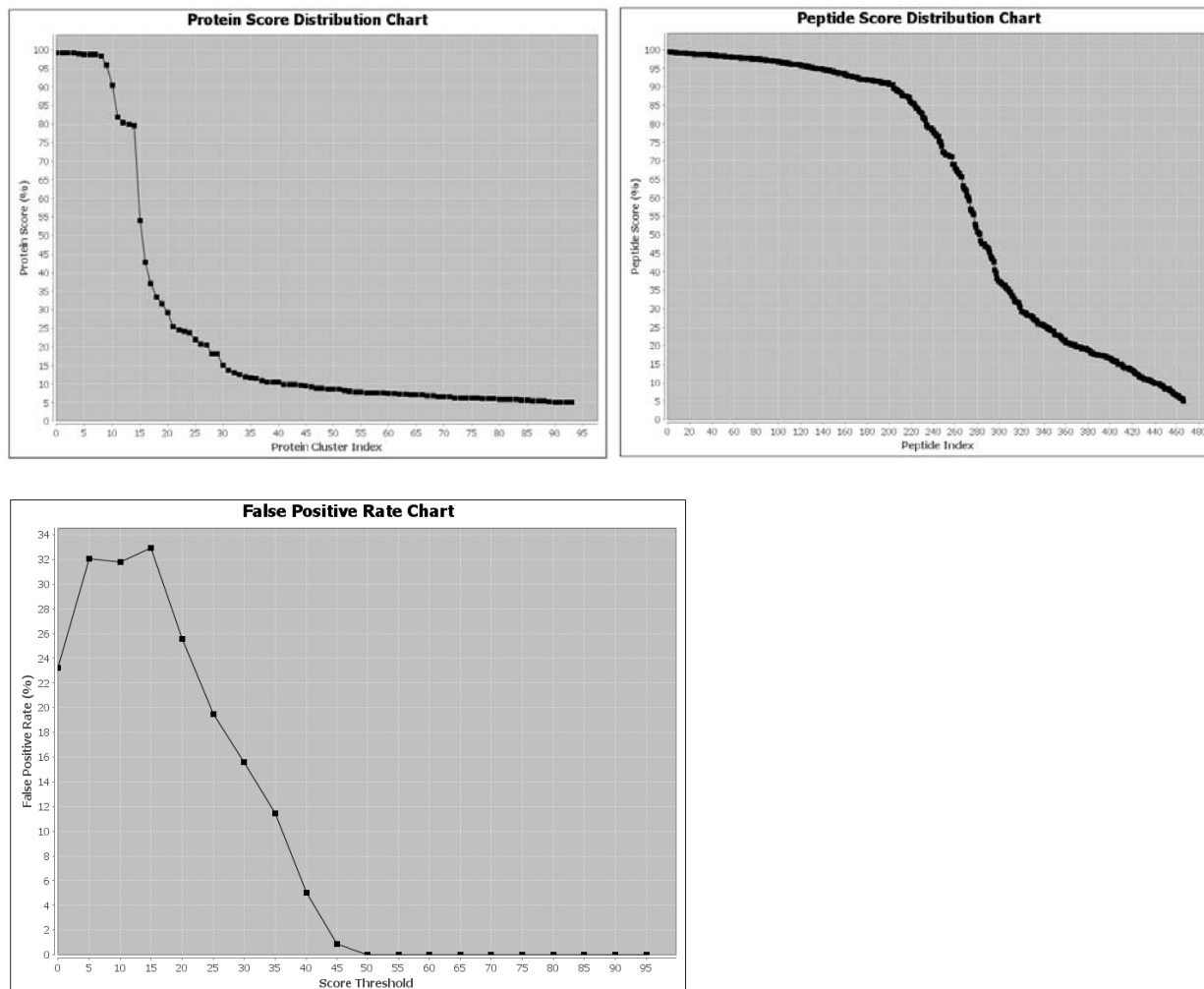
Select a Protein ID results file. Right click and select “Export Statistics Graph”. The following window will appear, giving you the option of exporting a “Protein Score Chart”, a “Peptide Score Chart” or a “False Positive Rate Chart:

The image shows a dialog box titled "Export Statistics Graph Specification". It has a blue title bar with a close button (X) in the top right corner. The dialog is divided into several sections:

- Statistics Graph Type:** This section contains four radio buttons: "De novo ALC Score Chart", "Protein Score Chart" (which is selected), "Peptide Score Chart", and "False Positive Rate Chart".
- Options:** This section contains three fields: "File Format:" with a dropdown menu set to "png", "Width:" with a numeric input set to "800", and "Height:" with a numeric input set to "600".
- Select Export File Destination:** This section contains a text field labeled "Image File:" and a "Browse ..." button.

At the bottom of the dialog, there are three buttons: "OK", "Cancel", and "Help".

Below are examples of the “Protein Score Chart”, the “Peptide Score Chart” and the “False Positive Rate Chart”, respectively.



Chapter 15

15. Advanced Configuration and Environment Preferences

15.1 PEAKS Environment Preferences

This section will describe the setting up the configuration of environmental preferences including general, instrument, search engine and ion editor configurations.

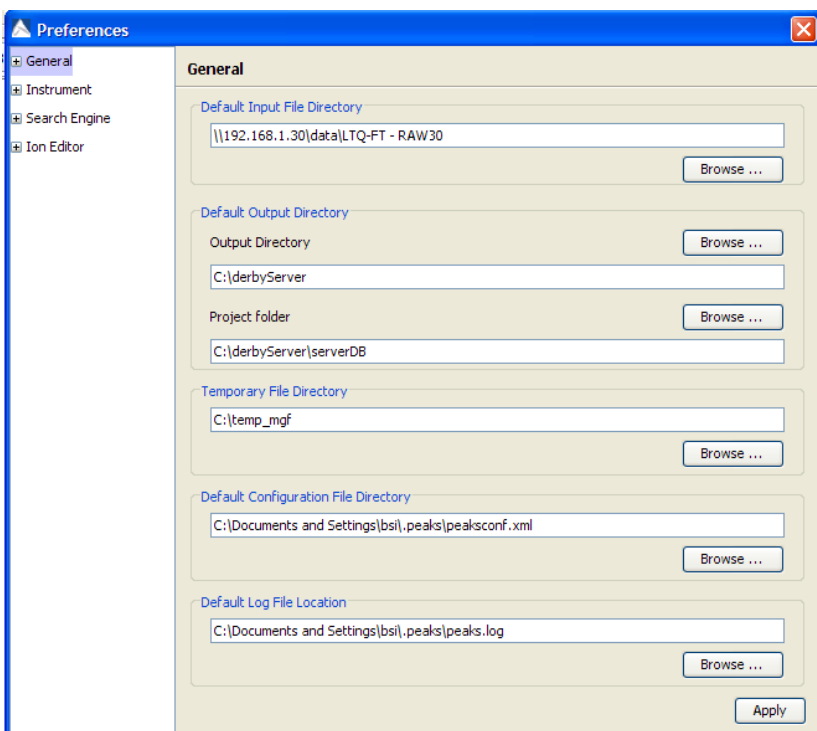
To begin click the Preferences toolbar icon 

Or

Select “Preferences” from the “Window” menu.

The following window will open:

Use the ‘ + ’ and ‘ – ’ boxes to expand and collapse the view.



15.1.1 General Preferences

Default Input File Directory- Select where your data is being inputted from using the “Browse” button.

Default Output File Directory- PEAKS outputs your results to C:\derbyServer by default. Select the “Browse” button to change this location.

Project folder- PEAKS uses C:\derbyServer\serverDB as the default project folder. Select the “Browse” button to change this location.

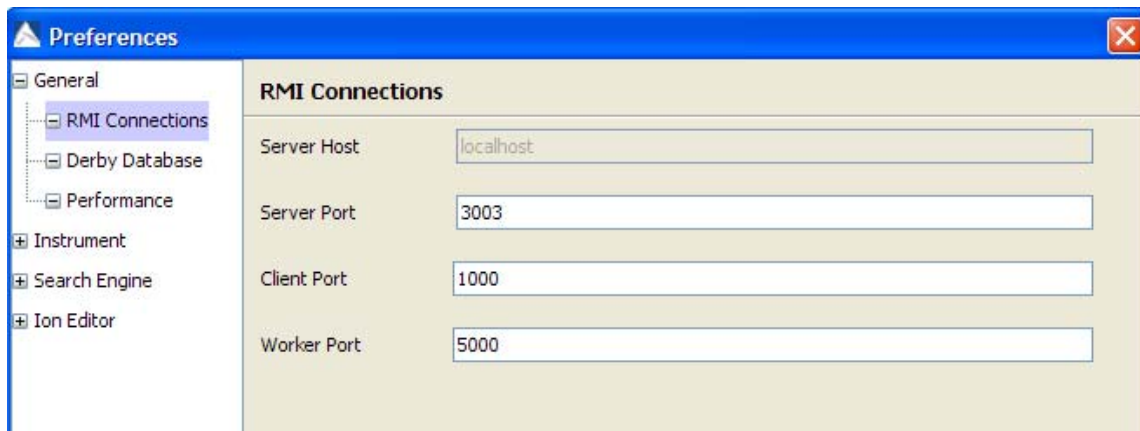
Temporary File Directory- PEAKS uses C:\temp_mgf as the default project folder. Select the “Browse” button to change this location.

Default Configuration File Directory- Your configuration files for PEAKS can be found at C:\Documents and Settings\bsi\peaks\peaksconf.xml by default. Select the “Browse” button to change this location.

Default Log File Location- Your log file for PEAKS can be found at C:\Documents and Settings\bsi\peaks\peaks.log by default. Select the “Browse” button to change this location.

RMI Connections

Clicking on “RMI connections” on the menu on the left hand will open the following window:

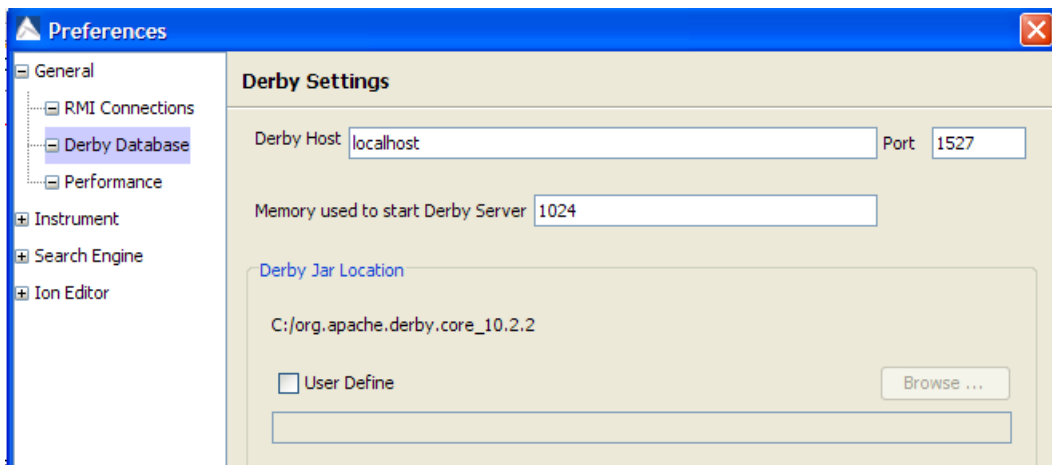


The default port numbers for the Server, Client and Worker will appear. The port numbers can be changed if conflicts arise. Contact technical support at BSI for more information.

Click the “Apply” button to save any changes that you made.

Derby Database

Clicking on “Derby Database” on the menu on the left hand will open the following window:



Derby Host

The name of the “Derby Host” as well as the “Port” number will come up by default and can be changed if needed.

Memory used to start Derby Server

The amount of “Memory used to start Derby Server” will also come up by default but can be changed if more memory is available.

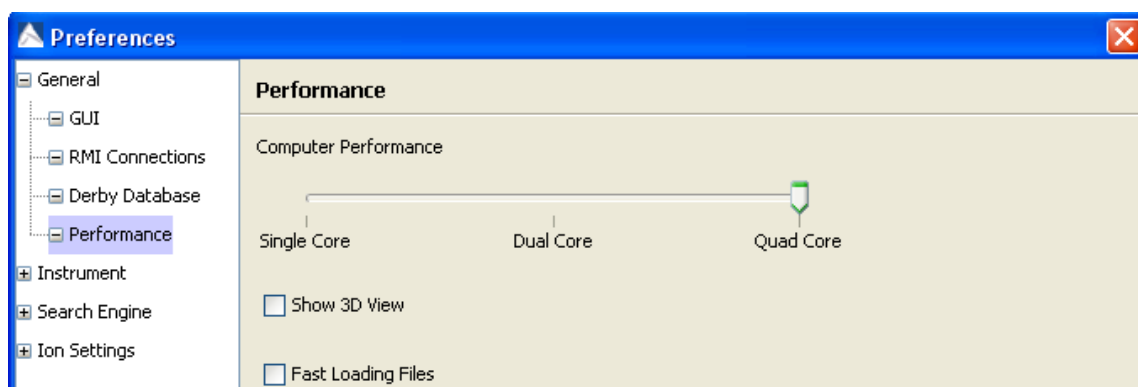
Derby Jar Location

The “Derby Jar Location” panel will list the location of the Derby Jar file by default. If you would like change this location, check the “User Define” box and click on the “Browse” button to select a new location.

Click the “Apply” button to save any changes you have made.

Performance

Clicking on “Performance” on the menu on the left hand will open the following window:



Computer Performance

Select the number of cores that your computer contains (i.e. single, double or quad core).

3D View

PEAKS will display a 3D view with your quantification results. Check the “Show 3D View” box to enable this function. PEAKS 5.1 comes with the Java3D program to support the viewing of 3D images.

Fast loading files

This function is for raw file loading. If you check the “Fast Loading Files” box, PEAKS will only load spectrum header information without the peak list. Please note that if you set up a project which contains more than 20 raw files, this function will not work well due to memory issues, and you should uncheck the “Fast Loading Files” box.

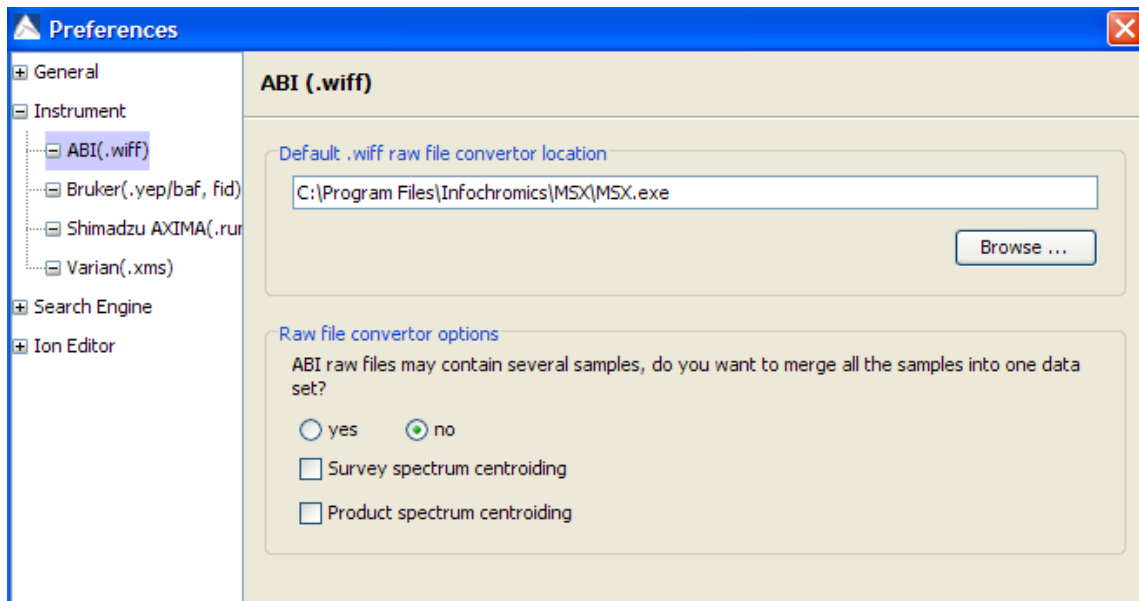
Click the “Apply” button to save any changes you have made.

15.1.2 Instrument Preferences

This section will allow you to change any preferences for the following instruments: ABI, Bruker, Shimadzu and Varian.

ABI (.wiff)

Clicking on “Instrument” and then “ABI (.wiff)” on the menu on the left hand side will open the following window:



Default .wiff raw file convertor location

Click “Browse” to tell PEAKS the location of the Default .wiff raw file converter.

Raw file converter options

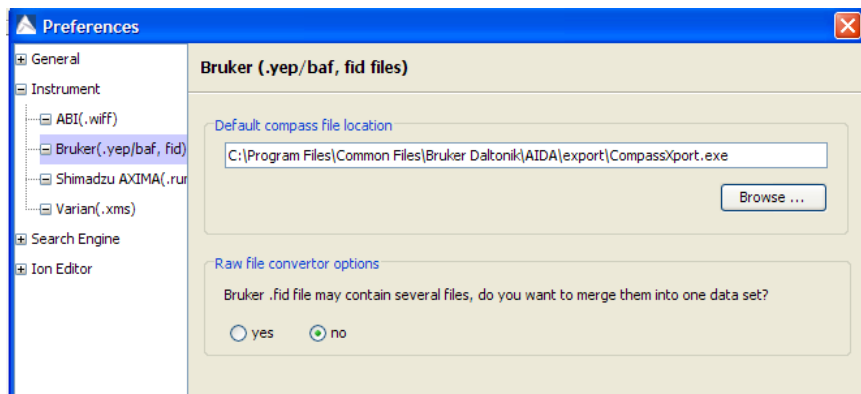
ABI raw files may contain several samples. By default, these samples are not merged into one data set. Select “yes” if you would like PEAKS to merge all the samples into one data set.

For PEAKS to work optimally, it is important to select if the survey spectrum or the product spectrum has been centroided.

Click the “Apply” button to save any changes you have made.

Bruker (.yep/baf, fid)

Clicking on “Instrument” and then “Bruker (.yep/baf, fid)” in the menu on the left hand side will open the following window:



Default compass file location

Click “Browse” to tell PEAKS the location of the CompassXport file converter.

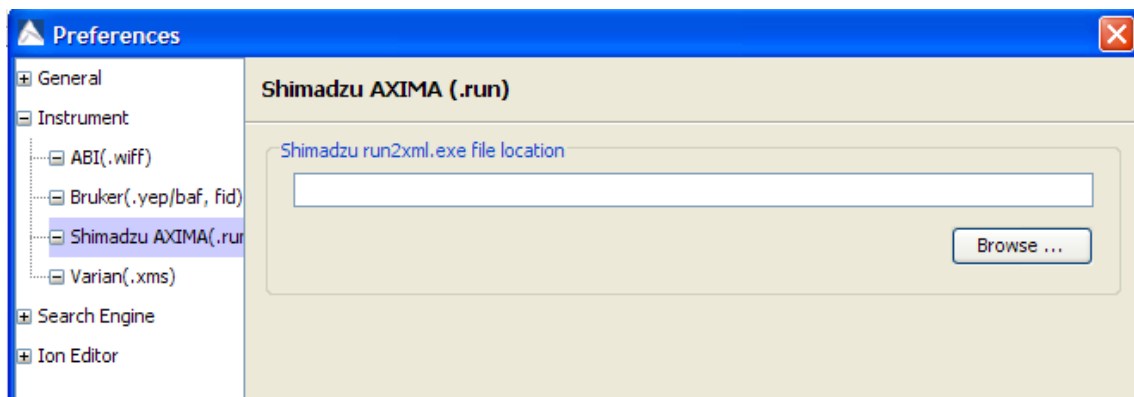
Raw file converter options

Bruker .fid files may contain several samples. By default, these samples are not merged into one data set. Select “yes” if you would like PEAKS to merge all the samples into one data set.

Click the “Apply” button to save any changes you have made.

Shimadzu AXIMA (.run)

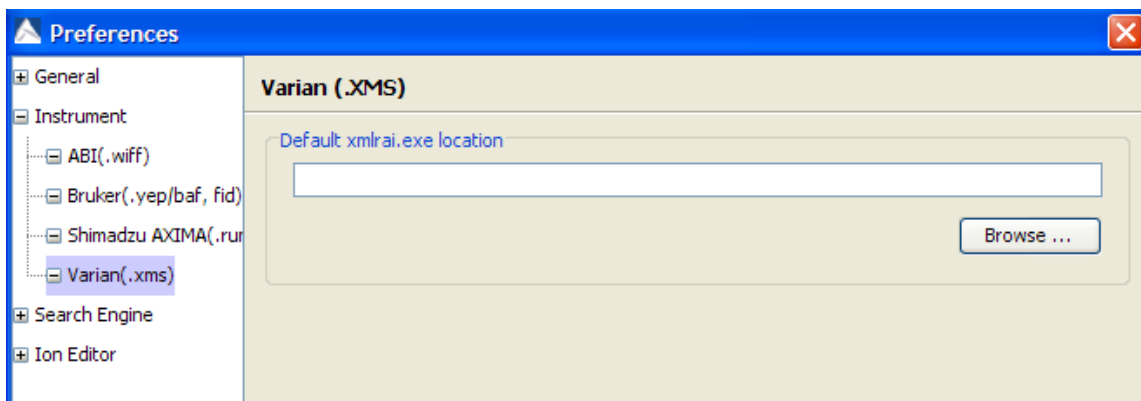
Clicking on “Instrument” and then “Shimadzu AXIMA (.run)” in the menu on the left hand side will open the following window:



Click “Browse” to tell PEAKS the location of the Shimadzu run2xml.exe file. Click the “Apply” button to save any changes you have made.

Varian (.xms)

Clicking on “Instrument” and then “Varian (.xms)” in the menu on the left hand side will open the following window:

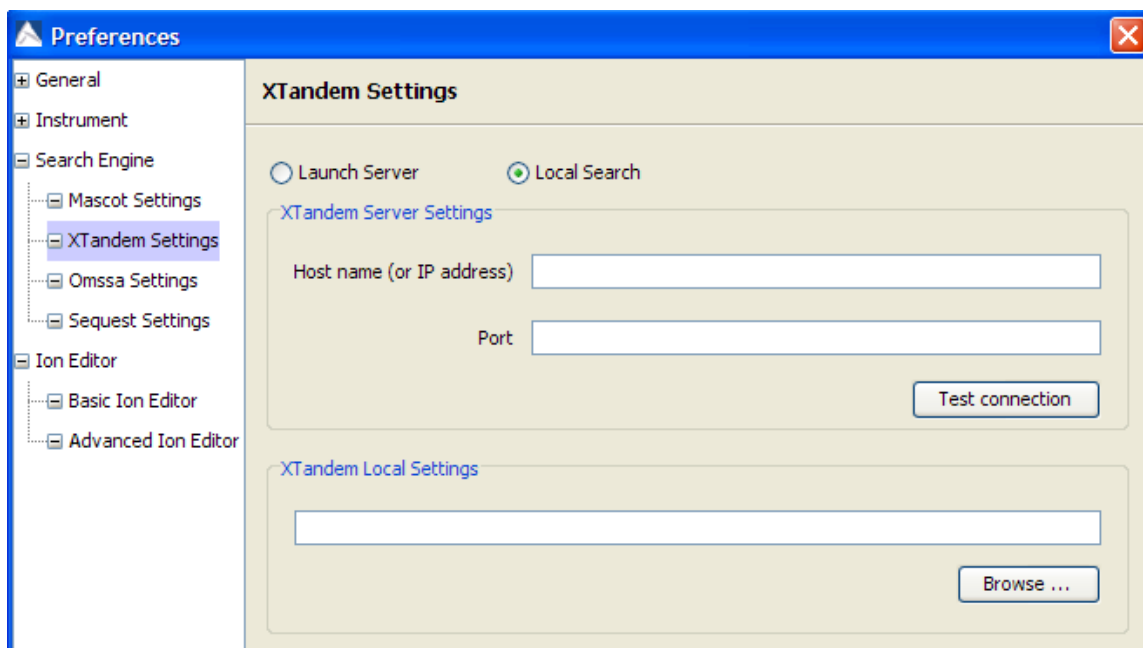


Click “Browse” to tell PEAKS the location of the xmlrai.exe file. Click the “Apply” button to save any changes you have made.

15.1.3 Search Engine Preferences

X!Tandem Settings

Clicking on “Search Engine” and then “X!Tandem Settings” in the menu on the left hand will open the following window:

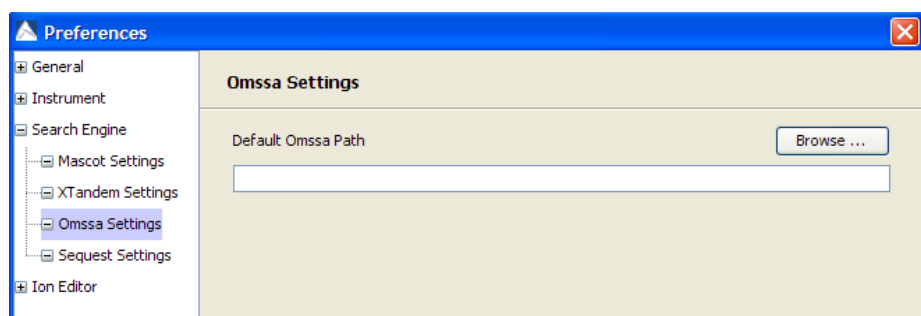


Start by selecting whether you would like PEAKS to access a server or local version of X!Tandem. If you select the server version, you must enter the Host name (or IP address) as well as the port. To make sure that you entered everything correctly and that the server is working, click the “Test Connection” button.

As PEAKS provides a local copy of X!Tandem, upon installation, a default path will appear in the Local Settings section. If you wish to use another license/location for X!Tandem, click the “Browse” button to tell PEAKS where to find the search engine. Click the “Apply” button to save any changes.

OMSSA Settings

Clicking on “Search Engine” and then “OMSSA Settings” in the menu on the left hand will open the following window:

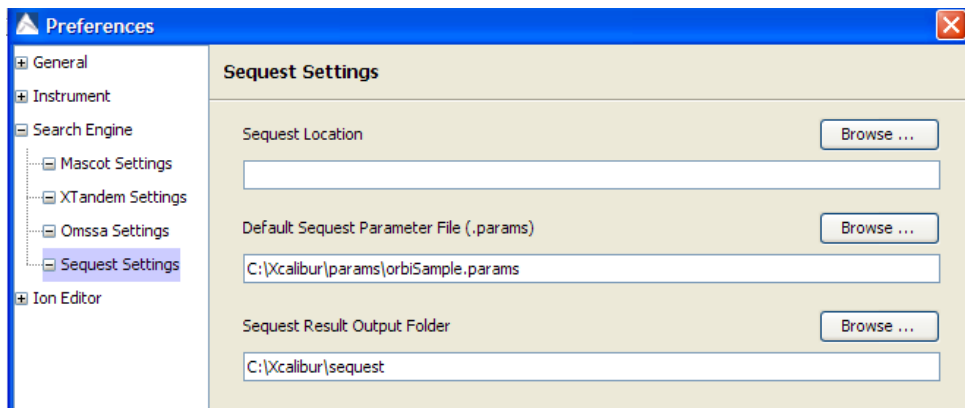


As PEAKS provides a local copy, upon installation, a default path will appear here. If you wish to use another license/location for OMSSA, click the “Browse” button to tell PEAKS where to find the desired path. Click the “Apply” button to save any changes you have made.

Sequest Settings

Clicking on “Search Engine” and then “Sequest Settings” in the menu on the left hand will open the following window:

To use Sequest click the “Browse” button to tell PEAKS where to find the search engine. You must also browse your computer to find the location of the “Default Sequest Parameter File

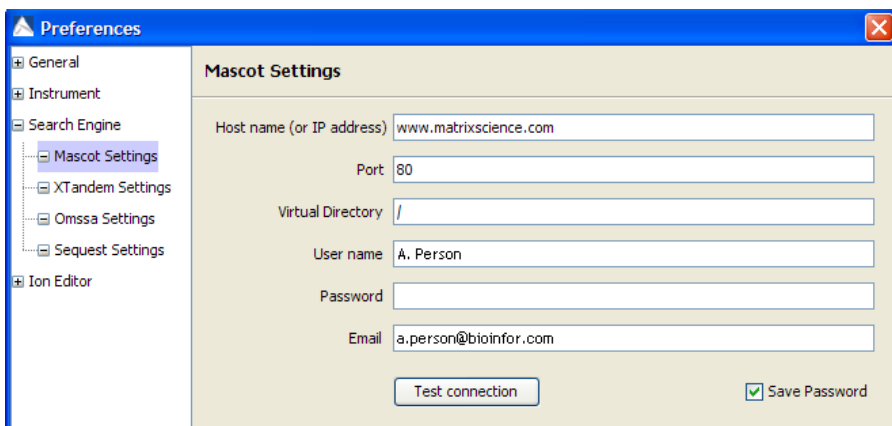


(.params)” as well as the “Sequest Result Output Folder”. Click the “Apply” button to save any changes you have made.

Mascot Settings

Clicking on “Search Engine” and then “Mascot Settings” on the menu on the left hand side will open the following window:

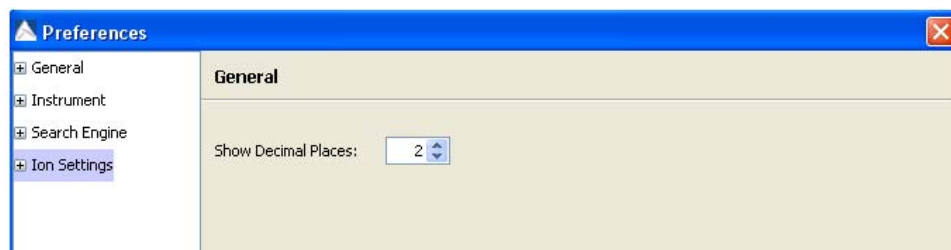
In this window you will tell PEAKS how to access your Mascot server (if applicable). Enter the Host name (or an IP address), Port, Virtual Directory as well as your user name, password and email address. To make sure that you entered



everything correctly and that the server is working, click the “Test Connection” button. If you would like to save your password so that you don’t have to enter it every time, check the “Save Password” box. Click the “Apply” button to save any changes you have made.

15.1.4 Ion Editor Preferences

Clicking on “Ion Editor” on the menu on the left hand will open the following window:



Decimal places

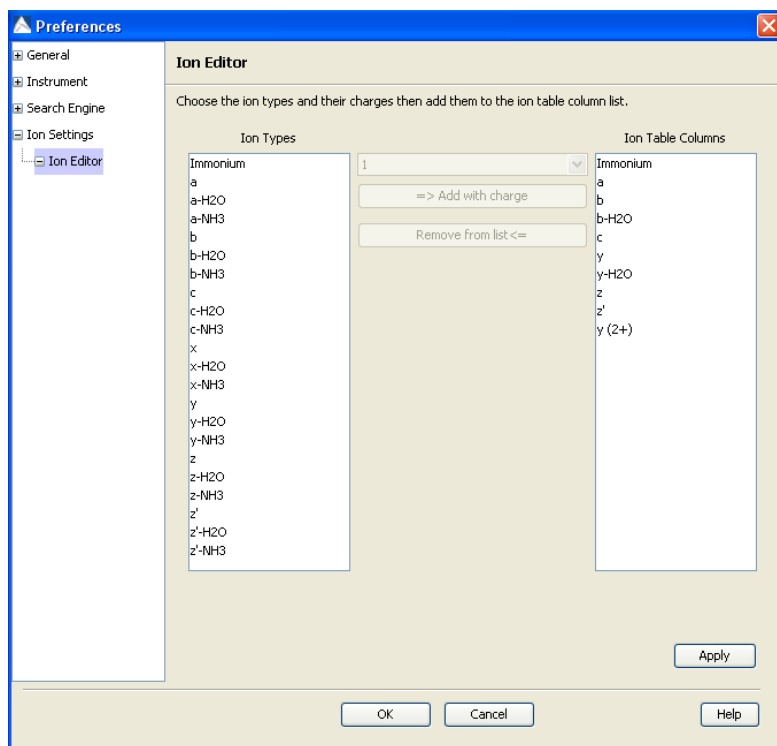
Select the number of decimal places you would like to appear in the ion table. The default is set to two decimal places.

Ion Editor

Clicking on “Ion Settings” and then “Ion Editor” in the menu on the left hand will open the following window:


To select an ion type to be viewed in the ion table, click on the ion type in the “Ion Type” list found on the left hand side of the window. You now need to select the charge for that ion type from the drop-down menu. Once you have done this, click on the “=>Add with charge” button and the ion type will now appear in the “Ion Table Columns” list on the right hand side of the window. To remove an ion type from the “Ion Table Column” list, select the ion type and click on the “Remove from list <=” button. The ion type will now appear in the “Ion types” list.

Click the “Apply” button to save any changes you have made.



15.2 PEAKS Configuration

This step includes configuration of enzymes, PTMs, databases, instruments, and parameters.

To begin click the Configuration toolbar icon 

Or

Select “Configuration” from the “Windows” menu.

15.2.1 Enzyme Configuration

PEAKS can use almost any enzyme, or combination of enzymes in your analysis. You can select from any of the built-in enzymes or define your own. From the “Configuration” window select “Enzyme” from the left hand to change your enzyme configuration.

Built-in enzymes

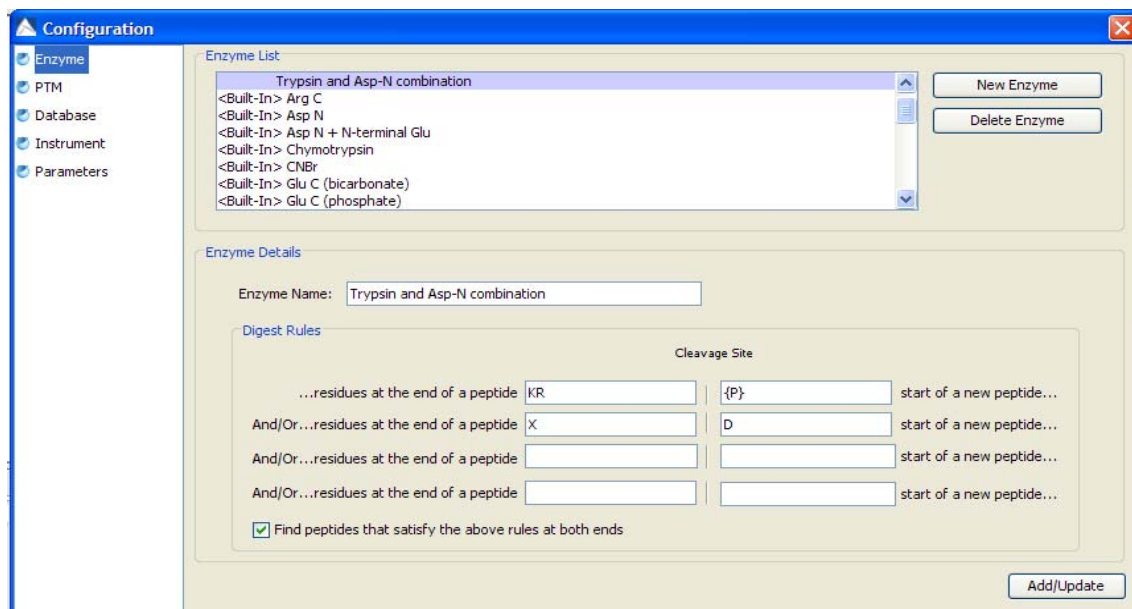
All of the built-in enzymes within PEAKS are listed in the “Enzyme list”. Clicking on one of these built-in enzymes will display the information listed about that enzyme in the “Enzyme Details” panel. Note that you cannot delete or change the details of a built-in enzyme and therefore the “Delete enzyme” button and the “Digest Rules” panel will be grayed out.

Create a new enzyme

Click on the “New Enzyme” button”.

Digest Rules: This is how you specify where you enzyme will cleave the protein between two amino acids to create peptides. The letter X denotes ‘any amino acid in this position’, while {set brackets} indicate any amino acid except the one in the brackets. You can also choose to select the check the box “Select peptides that satisfy the above rules at both ends” if you desire.

The example below shows a combination of Trypsin and Asp-N.



You must click the “Add/Update” button for the changes to be saved. Your new enzyme will now appear in the “Enzyme List” where you can access it later. If you wish to delete an enzyme that you created, select the appropriate enzyme and click the “Delete Enzyme” button.

Note: For information on defining new enzymes “on the fly” for PEAKS *de novo* or PEAKS Protein ID, see section [6.1.2 Enzyme Options](#) or [7.1.2 Enzyme Options](#), respectively.

15.2.2 PTM Configuration

From the “Configuration” window select “PTM” from the left hand to change your PTM configuration.

Built-in PTMs

The built-in PTMs within PEAKS are listed in the “PTM List”. To see additional built-in PTMs from the Unimod library, click the “Show unimod” box. Clicking on one of these built-in PTMs will display the information listed about that PTMs in the “PTM Details” panel. Note that you cannot delete or change the details of a built-in PTM and therefore the “Delete PTM” button and the “PTM Rules” panel will be grayed out.

Create a new PTM

Click on the “New PTM” button”. Now simply enter the information about your PTM in the “PTM Details” panel.

Name: this name will appear in the PTM list for future use after it is saved.

Monoisotopic mass: the mass that the residue gains or losses as a result of the PTM.

Neutral loss mass: the mass that the modified residue loses as a result of fragmentation. Ex. 28 would signify a loss of 28 Daltons.

Chemical formula: the chemical formula of the PTM. This should correspond to the mass listed above.

Residues that can be modified: Enter residues that can be modified anywhere, residues that can only be modified if they are at the N-or C-terminus or in the middle only.

Rule: you can enter a comment for your reference.

You must click the “Add/Update” button for the changes to be saved. Your new PTM will now appear in the “PTM List” where you can access it later. If you wish to delete a PTM that you created, select the appropriate PTM and click the “Delete PTM” button.

The screenshot shows the 'Configuration' window with the 'PTM' tab selected. On the left is a sidebar with 'Enzyme', 'PTM', 'Database', 'Instrument', and 'Parameters'. The main area is divided into two panels. The top panel, 'PTM List', shows a list of PTMs under the 'MyCarbo' category, with 'Oxidation on Methionine' selected. To the right of this list are 'New PTM' and 'Delete PTM' buttons. The bottom panel, 'PTM Details', contains fields for: PTM name (Oxidation on Methionine), Mass (Monoisotopic) (15.994915), Neutral loss mass (Monoisotopic) (0.0), Residues that can be modified (M, Anywhere), Formula (O), and Rule. An 'Add/Update' button is at the bottom right.

The example listed below is one where we knew that only methionine was oxidized.

Note: For information on defining new PTMs “on the fly” for PEAKS *de novo* or PEAKS Protein ID, see section [6.1.3 PTM Options](#) or [7.1.4 PTM Options](#), respectively.

15.2.3 Database Configuration

In addition to *de novo* sequencing of peptides, PEAKS 5.1 also has the ability to search through a database search to identify proteins. In order to use this function, PEAKS must have access to a protein or EST database in FASTA format or an EST database of DNA sequences. You can point PEAKS to an existing database on your system, or download one. Additionally, you can associate taxonomy with certain databases.

From the “Configuration” window select “Database” from the left hand to change your database configuration. The “Database list” at the top of the screen will show you databases that you have already configured. Select one of these files to see the details in the “Database Details” panel below.

Configure a new database

1) Select the “New Database” button on the right hand side of the “Database List”. You will now be filling in the specifics for your database in the “Database Details” panel below.

Database Details

FASTA format database: NCBI nr

Basic Options

Database name: NCBI nr Download Database

Path: C:\PEAKS database\NCBI nr Browse

☐ EST database

Advanced Options - Fasta Title Format

Rule to parse accession/id from FASTA title:

Rule to parse description from FASTA title:

Accession/id URL:
 <Accession/ID>

Delimiter:

Taxonomy Options

taxonid: C:\PEAKS database\NCBI nr\gi_taxid_prot.dmp.gz Browse Download

taxdmp: C:\PEAKS database\NCBI nr\taxdmp.zip Browse Download

2) *FASTA Format Database*: Select your database from this drop-down menu, or select “Other” if your database is not in the list or if you would like to submit your own database.

Database Details

FASTA format database: NCBI nr

Basic Options

Database name:

Path: C:\PEAKS database\NCBI nr Browse

☐ EST database

Download Database

3) In the basics option panel, enter a name for your database and select “Download Database”. The following window will appear:

Download options

?

You have choose to download the database from:

<ftp://ftp.ncbi.nih.gov/blast/db/FASTA/nr.gz>

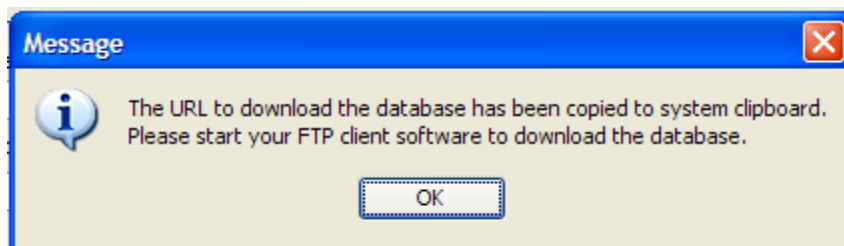
Click **YES** to invoke your default FTP client software and download automatically. It may fails due to the setting conflict of default FTP client software.

Click **NO** to copy this URL to system clipboard. You can start your FTP client software manually.

After download the database, you need do decompression if it's a compressed file. And then put the database file in the directory that PEAKS can access.

Yes No

4) If you would like to invoke your default FTP client software and download automatically, click “Yes”. If you select “No” the following window will appear telling you that the URL will be copied to your system clipboard. Click “Ok”.



Open your Internet Explorer and paste the URL into the address bar. A file download window will open. Click Save.

5) Once the database is downloaded, you need to make sure that you decompress the file if it is compressed using a program such as WinZip, or WinRar to extract its contents. The file inside the compressed file will be a FASTA format text file (a .fas or a .fasta file).

6) Finally put the database file into a directory that PEAKS can access.

7) Click “Browse” to tell PEAKS where the database file is located.

8) If the database that you have selected is an EST database, check the box labeled “EST Database”. If not, leave it blank.

9) Since you have already selected a “FASTA Format Database” in Step 2, the Accession number information and the parsing rules for the database headers are shown in the textboxes below in the “Advanced Options- Fasta Title Format” panel.

If you chose “Other” (in step 2) you must enter parsing parameters by completing the textboxes. Alternatively, if the database format is the same as one of the public databases, you can choose to apply that database’s format when PEAKS reads the database. Select the database that is similar to yours from the drop-down list to fill the textboxes with the appropriate parsing rules.

A note on parsing rules

Apart from starting with a “greater than” symbol, the precise syntax of the FASTA title line varies from database to database. For this reason, PEAKS uses Java *Regular Expressions* to

define how the accession string and the description text should be parsed from the FASTA title line.

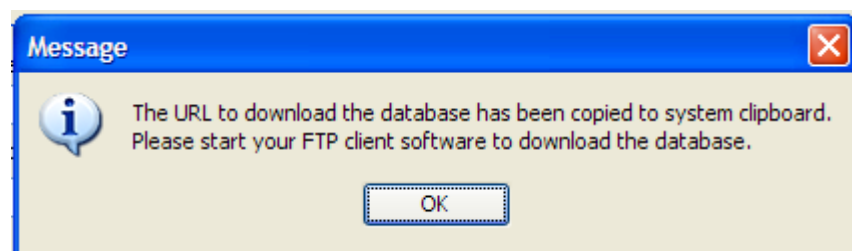
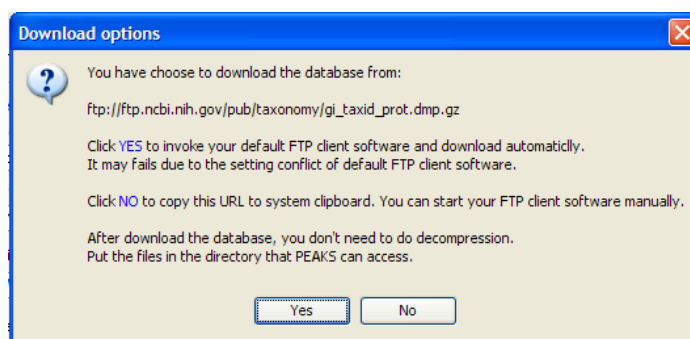
A note on using a delimiter

Some databases use one entry to represent multiple protein entries. The FASTA headers are concatenated with a delimiter. Since some of these databases use unprintable control codes as delimiters PEAKS will use the equivalent ASCII decimal code to represent them. For example the NCBI NR database uses CTRL-A as a delimiter so the user should input "1" as its equivalent decimal delimiter as listed here.

10) To be able to do PEAKS Protein ID using a specific taxonomy, you will need to download some files and place tell PEAKS where to find them in the "Taxonomy Options" panel.

11) To download the taxonid file, click the "Download" button. The following window will appear:

If you would like to invoke your default FTP client software and download automatically, click "Yes". If you select "No" the following window will appear telling you that the URL will be copied to your system clipboard. Click "Ok".



Open your Internet Explorer and paste the URL into the address bar. A file download window will open. Click Save. Repeat Step 11 taxdmp file. Please not that you do not have to

decompress the taxonomy files.

12) Now that you have downloaded the taxonomy files, you must tell PEAKS where to access them by clicking the Browse button and selecting the file.

13) To save the database to your "Database List", you must click the "Add/Update" button before clicking the "Ok" button.

Delete a previously saved database

If you would like to delete a database file, select the database that you wish to delete and click on the "Delete Database" button.

Set a database as default

Select the file and click the set as default button which is located to the right of the "Database List". This database will now be used by PEAKS when you run PEAKS Protein ID.

Moving/Updating a Database

If you choose to move a database to another directory or delete it entirely, you need to notify PEAKS. You must remove the database from the list and re-load it. Until you do so, the database name will appear in red in the list of databases and any protein identification using that database will fail.

If you choose to update the database (perhaps by downloading the latest database file and overwriting the old database file), PEAKS will show the database information in light gray. A *light grey color could also mean that the database does not have header information.*

Best practices: configuring databases for use with X!Tandem

At the time of this writing, X!Tandem had trouble searching through large databases, and would crash. It is therefore suggested that X!Tandem only be used with small databases; or if used with a large database, a taxon should be specified. The NCBI nr and SwissProt databases are ideal for this purpose.

Best practices: configuring databases for use with OMSSA

At the time of this writing, we could not use OMSSA with databases that were not in NCBI format, or SwissProt format, and have those results available to inChorus.

Also, a bug in OMSSA prevents us from easily using databases with OMSSA when they are stored in a folder that contains a space in its path. This creates problems when PEAKS creates temporary databases on our behalf. To avoid this, best practices suggest that you put all our databases in a folder “c:\peaksdatabases”. The folder “c:\my documents\databases” wouldn’t work because it contains a space between ‘my’ and ‘documents’. Using spaces in the database file name causes the same problem. So after you download and extract our database you should call the database file “ncbinr.fas”, or “ncbi_nr.fas” rather than “ncbi nr.fas”.

15.2.4 Instrument Configuration

From the “Configuration” window select “Instrument” from the left hand to change your instrument configuration.

Built- in Instruments

Select the manufacturer of your instrument from the drop-down list. The names of the instruments will then appear in their vendor specific formats. Select your instrument and you will be able to view the information on your instrument in the “Instrument details” panel below. You can also select “General” in the manufacturer list and the instruments will be listed in a general format.

Note that you cannot delete or change the details of a built-in instrument and therefore the “Delete PTM” button and the “Instrument Details” panel will be grayed out.

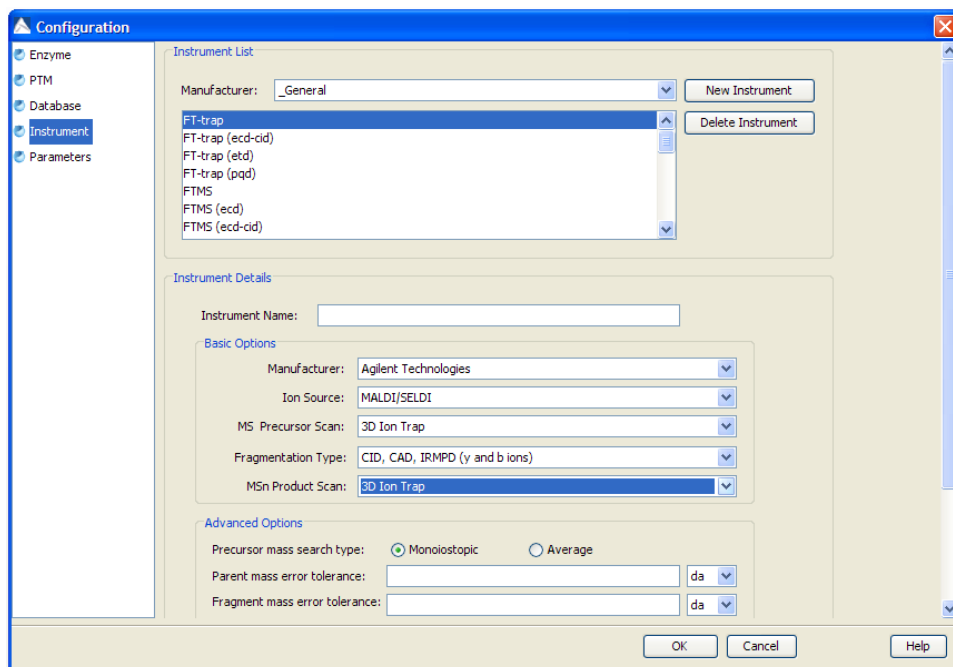
Create a new instrument

Click on the “New Instrument” button and the following window will appear:

In the “Instrument Details” panel, create a name for your instrument.

Next, fill in your details in the “Basic Options” panel. In the manufacturer drop-down list, select a specific vendor or “General”.

Ion Source: Use the drop-down list to select what ion source that was used; MALDI/SELDI or ESI(nano-spray). This will help the PEAKS Data Refine tool to decide the charge of the ions.



MS- Precursor Scan: Use the drop-down list to select what type of MS scan was performed. This selection will tell the PEAKS Data Refine tool if the survey scan is of sufficient resolution to determine the charge and the monoisotopic peak from the examination of the survey scan.

Fragmentation type: Use the drop-down list to select the method of fragmentation that was used. This selection will tell PEAKS what type of ion-series to expect for PEAKS auto *de novo* sequencing and PEAKS protein ID database search. Select CID/ECD if alternating fragmentation is used to allow the algorithm to determine the type of fragmentation from each scan header.

MSⁿ Product Scan: Use the drop-down list to select what type of MSⁿ scan was performed. This selection will help PEAKS decide which internal parameters (for weighing fragments and amount of noise) to use during PEAKS auto *de novo* sequencing and PEAKS protein ID database search. Select LIT/FT if alternating hi-res/low-res modes are used, allowing the algorithm to determine the mass analyzer from the scan header. You can also use the “Advanced Options” to specify additional parameters.

Precursor Mass Search Type: Select “Monoisotopic” or “Average. For ion-trap instruments, it is usually beneficial to allow the PEAKS protein ID database search to use an average mass.

Parent and Fragment error mass tolerance: User specified values. These will appear on the PEAKS *de novo* and PEAKS protein ID options screens when the instrument is selected.

Target Ions: Select which ions that you would like PEAKS *de novo* and Protein ID to focus their search on.

You must click the “Add/Update button for the changes to be saved. Your new instrument will now appear in the “Instrument List” where you can access it later. If you wish to delete an instrument that you created, select the appropriate instrument and click the “Delete Instrument” button.

15.2.5 Parameter Configuration

From the “Configuration” window select “Parameters” from the left hand to change your parameter configurations. Please note that you can only view and delete parameters from within this parameter window. From the “Parameter type” drop-down list at the top of the screen you can select *de novo*, PEAKS Parameters or SPIDER Parameters. The parameters that you have saved within these categories will be displayed below in the list. Select the parameter file that you would like to view.

Creating a new parameter file

If you would like to create and save new parameters you can do this when/before you set up auto *de novo* sequencing (see section [6.1 Setting up Auto *de novo* Sequencing Parameters](#)), PEAKS protein ID (see section [7.1 Setting up Protein Identification Parameters](#)) or SPIDER (see section [8.1 Setting up SPIDER Parameters](#)). These references will provide you with an explanation of all of the parameters.

Deleting a previously saved parameter file

If you would like to delete a parameter file, select the file that you wish to delete and click on the “Delete” button.

Viewing a previously saved parameter file

Selecting a file will display the details of that file below. For an explanation of the parameters, please see the pages listed in the “Creating a new parameter file” section above.

16. PEAKS Quantification Module



Many approaches to protein quantification using mass spectrometry data have been described in the literature. In terms of their implementation, most of them can be classified into three protocols.

- **MS:** Quantification based on the relative intensities of extracted ion chromatograms (XICs) for precursors within a single data set. This is the most widely used approach, which can be used with any chemistry that creates a precursor mass shift. For example, ICAT and SILAC.
- **MS/MS:** Quantification based on the relative intensities of fragment peaks at fixed m/z values within an MS/MS spectrum. For example, iTRAQ and Tandem Mass Tags.
- **Label free:** Label free quantification based on the relative intensities of extracted ion chromatograms (XICs) for precursors in multiple data sets aligned using mass and elution time.

All three protocols are fully implemented within PEAKS Q. The flow chart is shown below:



16.1 Setting up PEAKS Q Parameters

- 1) In the “Project View Frame”, select a  PEAKS Search result file.
 - 2) Click the PEAKS quantification toolbar icon .
- Or
- Select “Quantification” from the “Tools” menu.

The following window will open, displaying the quantification parameters:

Quantification

Tools

- . Data Refine
- . De novo
- . PEAKS Search
- . SPIDER Search
- . PTM Finder
- ✓ Quantification

Quantification

Basic Options

Mass Error Tolerance: Da Upper Bound of Precursor Charge:

Retention Time Range: min.

Label Options

☒ Labelling occurs at the MS/MS level eg. iTRAQ

☐ Labelling occurs at the MS level eg. ICAT

☐ Label-Free

Sample	Reporter Ion (Da)	Labelling Efficiency (%)
--------	-------------------	--------------------------

Slightly different options will be available if you select labeling at the MS level:

Quantification

Tools

- . Data Refine
- . De novo
- . PEAKS Search
- . SPIDER Search
- . PTM Finder
- ✓ Quantification

Quantification

Basic Options

Mass Error Tolerance: Da Upper Bound of Precursor Charge:

Retention Time Range: min.

Label Options

☐ Labelling occurs at the MS/MS level eg. iTRAQ

☒ Labelling occurs at the MS level eg. ICAT

☐ Label-Free

Sample	Added Mass	Residues	Labelling Efficiency (...)
--------	------------	----------	----------------------------

Quantification parameter options include the following:

Basic Options

Mass Error Tolerance: Quantification is based on the feature of a peptide that identifies its origin in the sample mixture. For example, in a SILAC experiment, one feature is unmodified peptides and the other is peptides modified with Label:13C(6) on Arginine or Lysine. For iTRAQ, the feature would be reporter ion m/z value. The mass error tolerance is for pairing up features.

Upper Bound of Precursor Charge: The peptide may present in different charges. Upper bound of precursor charge defines the maximum charge of peptides which are used for counting quantity.

Retention Time Range: The retention time range is for pairing up features. For iTRAQ, it is optional.

Labeling Options

Labeling occurs at the MS/MS level e.g. ITRAQ: It is for quantification based on the relative intensities of fragment peaks at fixed m/z values within an MS/MS spectrum.

Labeling occurs at the M level e.g. ICAT: It is for quantification based on the relative intensities of extracted ion chromatograms (XICs) for precursors within a single data set.

Label-free: It is for quantification based on the relative intensities of extracted ion chromatograms (XICs) for precursors in multiple data sets aligned using mass and elution time.

Sample: It is for specifying sample name.

Reporter ion: It is for specifying mass of reporter ion

Added Mass: The modified mass of a residue.

Residues: The residue to be modified.


Labeling efficiency: It is for specifying efficiency of chemical reaction.

Add label: It is used to add a label.

Delete label: It is used to delete a label.

16.2 3D View

In order to produce a 3D view, you must first select this in your preferences.

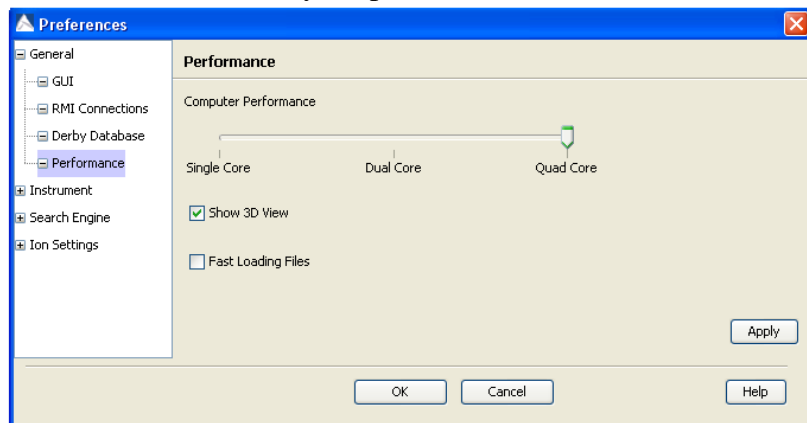
1) Click on the Preferences toolbar icon .

or select “Preferences” from the “Windows” menu.

2) Select “General” from the panel on the left hand side.

3) Select “Performance” and check the “Show 3D View” box.

4) Click “Apply”.



When a PEAKS Q run is complete, a “3D View” tab can be found in the “Peptide View”

window:

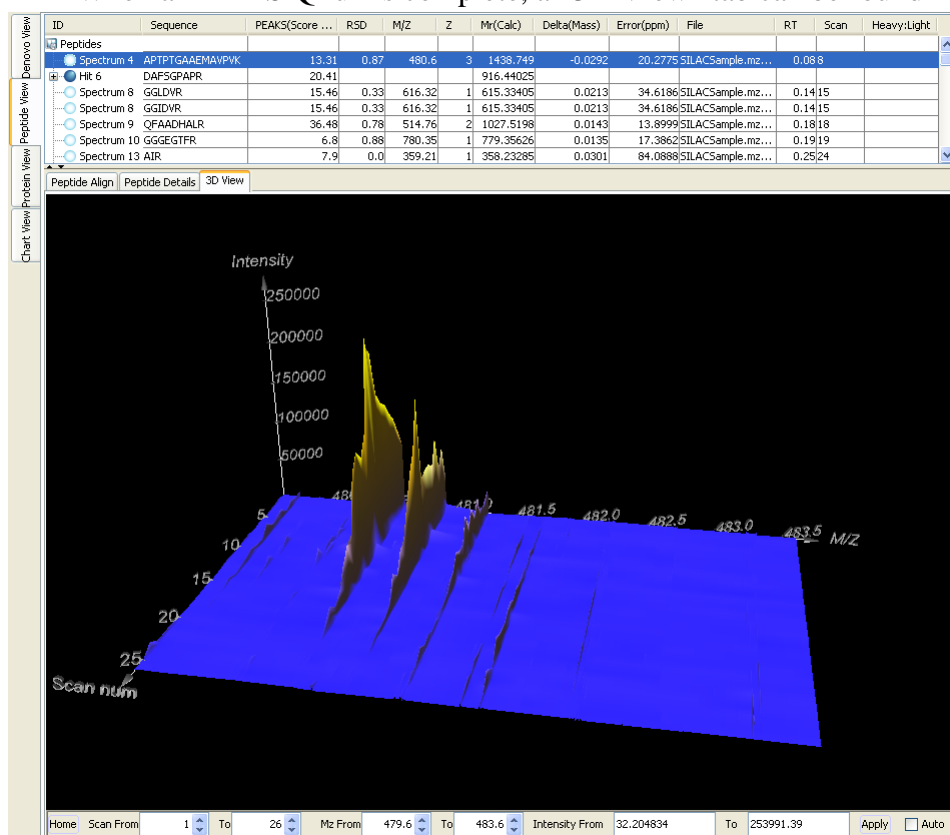
The 3D View contains 3 axes: intensity, m/z ratio and scan number.

The panel along the bottom allows you to narrow in on the peptides that you would like to examine.

You can specify a particular scan number range, m/z range or intensity range.

Click the “Apply” button to change the 3D view to your specified values.

Check the “Auto” button” and click “Apply” to the default numbers.




16.3 iTRAQ Walkthrough

Isobaric Tagging for Relative and Absolute Quantification (iTRAQ) uses isotopic labeling to enable relative quantitative comparisons. Up to eight different proteomic samples can be labeled using eight different isobaric tags.

16.3.1 Setting up a Database

Before running the walkthrough data, you need to set up a database. So that this can be a quick process, we have provided you with a sample fasta database called “SampleDB.fasta” in your PEAKS program folder (C:\PEAKS Studio 5.1\Data).

Click on the configuration toolbar icon  or select “Configuration” from the “Tools” menu. Select “Database” from the left hand side of the window. Under “Database Details” enter the following information:

Database Details

FASTA format database:

Basic Options


Database name:

Path:

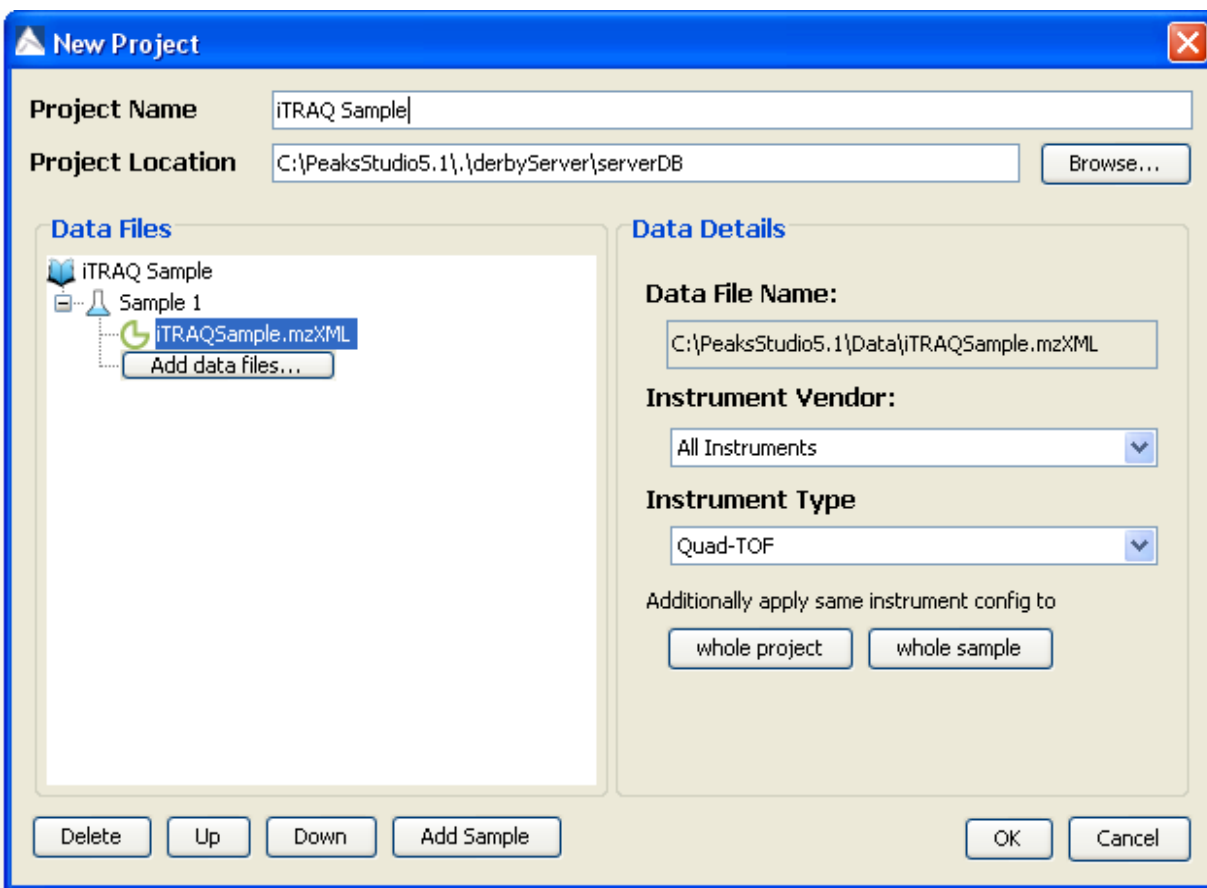
☐ EST database

You do not need to change any of the other information listed. Click the “Add/Update” button and then click “OK”.

16.3.2 Creating a Project

Click on the “Create new project” icon  or select “New project” from the “File” menu. Give a name to your project, for example “iTRAQ Sample”. Click Add data files. Locate “iTRAQSample.mzxml” from C:\PeaksStudio5.1\Data.

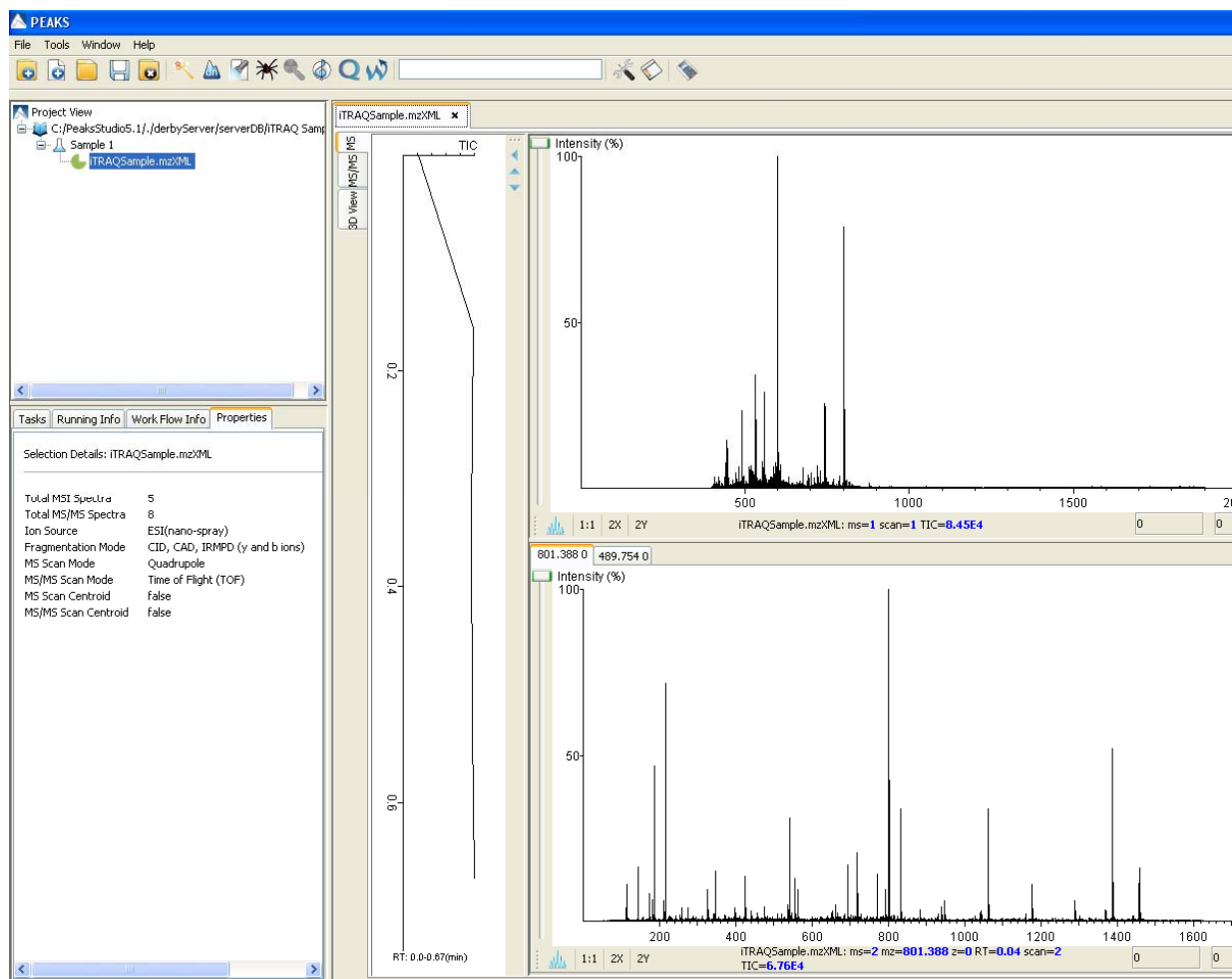
With the file highlighted, select, Quad-TOF from the “Instrument Type” drop-down list.



Click “OK”.

Double click on the iTRAQ sample file.

Your “Main Processing Screen” should look something like this:



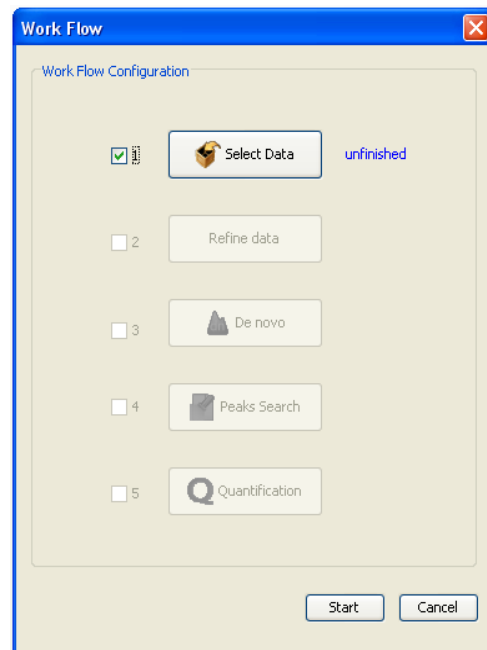
For ease of setup, this walkthrough will use the workflow tool in order to set all of the parameters at once. You will need to wait until the file has loaded (i.e. the file node looks like this:).

See section [16.2 3D View](#) to enable 3D view.

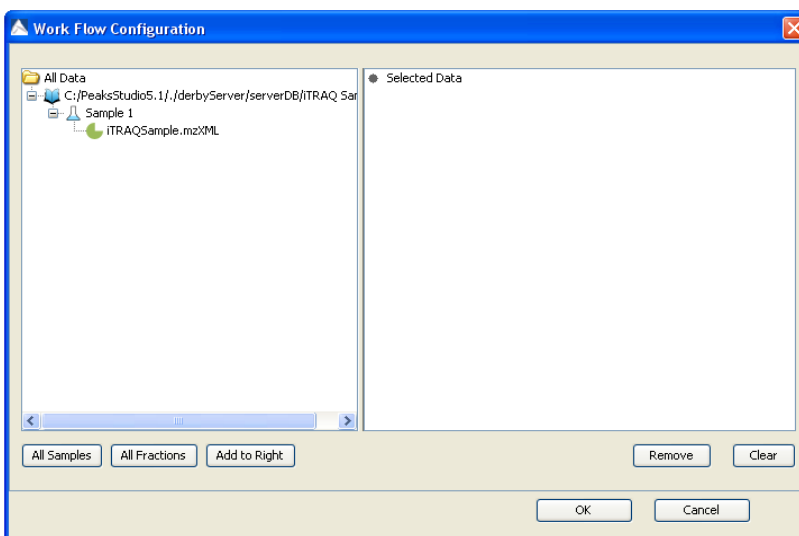
16.3.3 Setting up the Workflow Parameters

Click on the workflow button “”. The following window will appear:

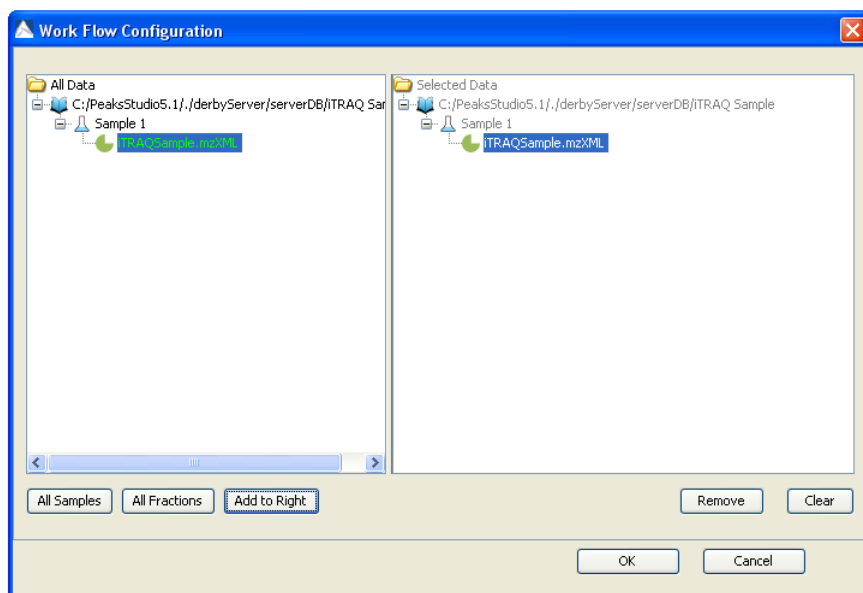
Click “Select Data”.



The following window will open:

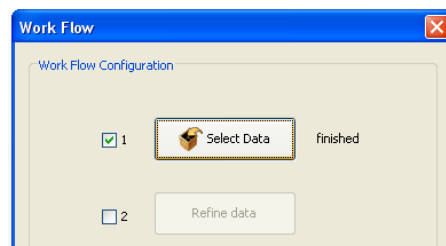


Select the iTRAQSample.mzXML file and then click “Add to Right”. The samples should now appear in the “Selected Data” panel as shown below:



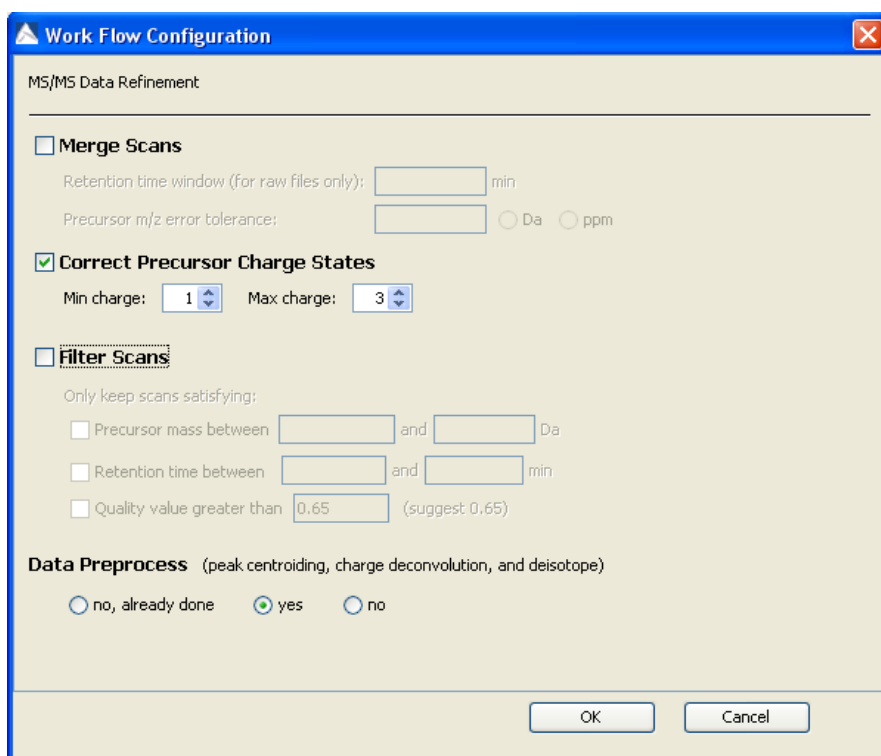
Click “OK”.

The workflow window should now display “finished” beside the “Selected Data” button.



Check the 2nd box and then click the “Refine data” button.

Enter the following parameters:



Work Flow Configuration

MS/MS Data Refinement

☐ **Merge Scans**

Retention time window (for raw files only): min

Precursor m/z error tolerance: Da ☐ ppm

☒ **Correct Precursor Charge States**

Min charge: Max charge:

☐ **Filter Scans**

Only keep scans satisfying:

☐ Precursor mass between and Da

☐ Retention time between and min

☐ Quality value greater than (suggest 0.65)

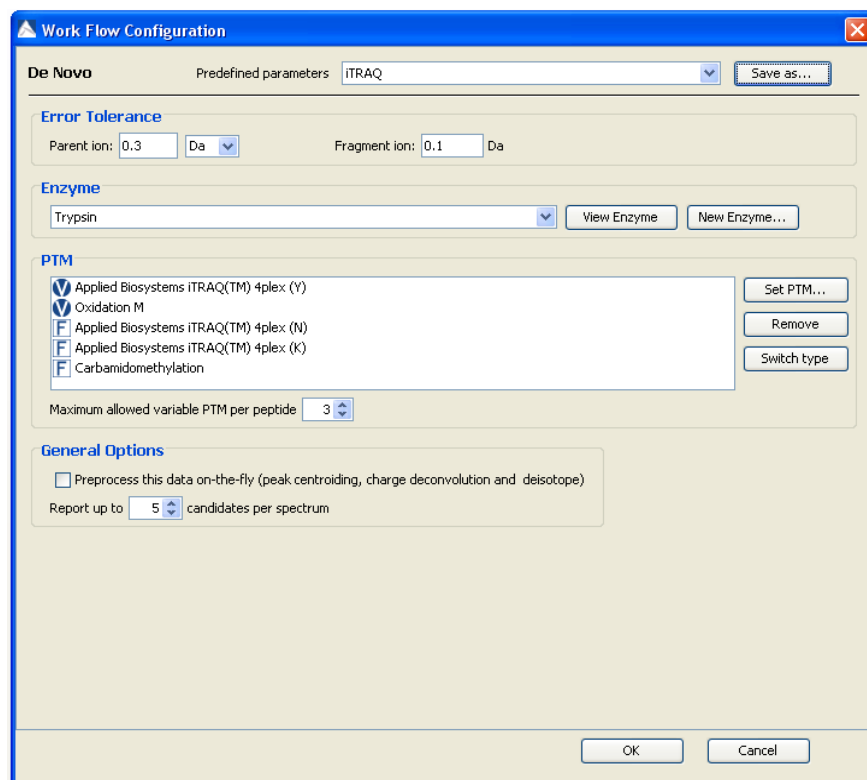
Data Preprocess (peak centroiding, charge deconvolution, and deisotope)

☐ no, already done ☒ yes ☐ no

OK Cancel

Click “OK”.

Check the 3rd box and then click the “*de novo*” button. Enter the following parameters:



Work Flow Configuration

De Novo Predefined parameters: Save as...

Error Tolerance

Parent ion: Da Fragment ion: Da

Enzyme

View Enzyme New Enzyme...

PTM

☒ Applied Biosystems iTRAQ(TM) 4plex (Y)
☒ Oxidation M
☐ Applied Biosystems iTRAQ(TM) 4plex (N)
☐ Applied Biosystems iTRAQ(TM) 4plex (K)
☐ Carbamidomethylation

Set PTM...
Remove
Switch type

Maximum allowed variable PTM per peptide:

General Options

☐ Preprocess this data on-the-fly (peak centroiding, charge deconvolution and deisotope)

Report up to candidates per spectrum

OK Cancel

Enter the PTMs as shown below by selecting “Set PTM”:

PTM Options

All PTM

Name	Mono mass	Residue site
N-acyl diglyceride cysteine	788.7258	[C]
N-isopropylcarboxamidomethyl	99.0684	[C]
N-SuccinimidyI-3-morpholine acetate	127.0633	[K], [X]@N
Oxidation M	15.9949	[M]
Oxidation HW	15.9949	[HW]
Palmitoylation	238.2297	[CSTK]
Phosphopantetheine	340.0858	[S]
Phosphorylation	79.9663	[YHCDR]
Phosphorylation-ST	79.9663	[ST]
Pyridoxal phosphate	229.0140	[K]
Pyro-glu from E	-18.0106	[E]@N
Pyro-glu from Q	-17.0265	[Q]@N
Pyrrolidone carboxylic acid	-17.0266	[Q]@N
S-carbamoylmethylcysteine cyclization(N-te...)	-17.0265	[C]@N
Sodium adduct	21.9819	[DE], [X]@C
S-pyridylethylation	105.0578	[C]
Sulfation	79.9568	[Y]

☐ Show unimod New...

Selected Fixed PTM

- Applied Biosystems iTRAQ(TM) 4plex (N)
- Applied Biosystems iTRAQ(TM) 4plex (K)
- Carbamidomethylation

Selected Variable PTM

- Applied Biosystems iTRAQ(TM) 4plex (Y)
- Oxidation M

Deselect Clear All Switch Type OK Cancel

Click “OK”. Check the 4th box and then click the “PEAKS Search” button.

Enter the foll

Work Flow Configuration

Database Search Predefined parameters: iTRAQ Save as...

Error Tolerance

Parent ion: 0.3 da using: monoisotopic mass Fragment ion: 0.1 Da

Enzyme

Trypsin View Enzyme New Enzyme...

Maximum missed cleavages per peptide: 1

PTM

- ☒ Applied Biosystems iTRAQ(TM) 4plex (Y)
- ☒ Oxidation M
- ☐ Applied Biosystems iTRAQ(TM) 4plex (N)
- ☐ Applied Biosystems iTRAQ(TM) 4plex (K)
- ☐ Carbamidomethylation

Set PTM... Remove Switch type

Maximum allowed variable PTM per peptide: 3

Database

☒ Select database Database: Sample DB New... Edit...

☐ Paste sequence Taxa: all species Set/View taxa...

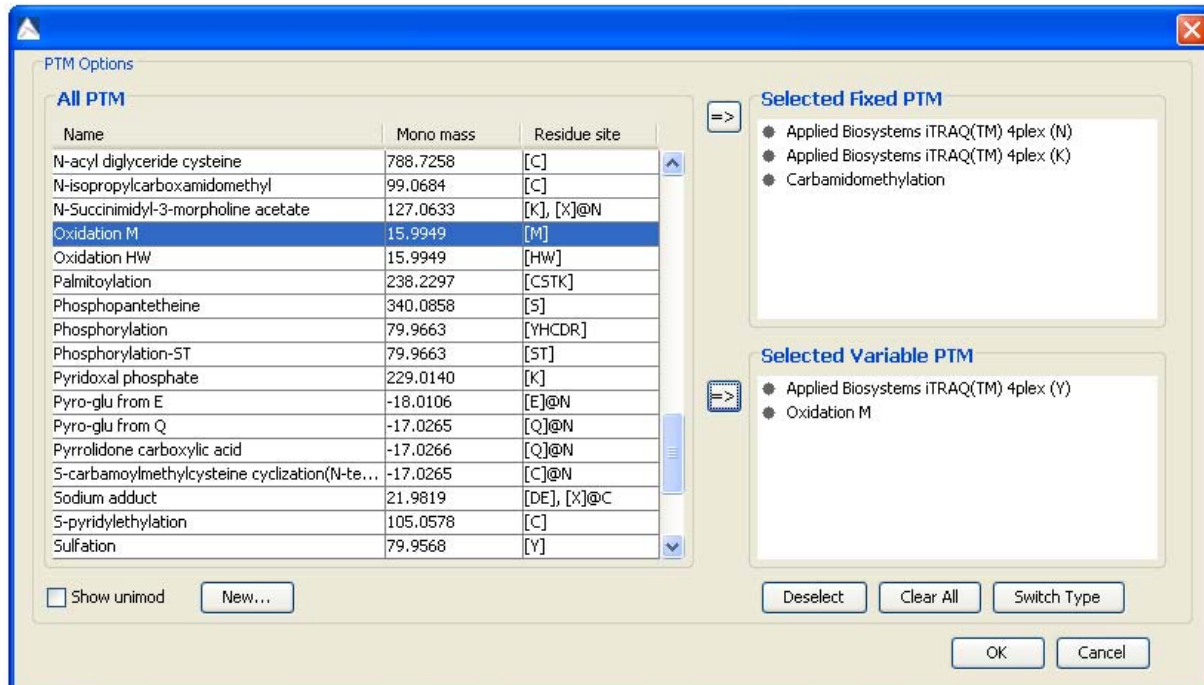
General Options

☐ Validation with reverse database as decoy

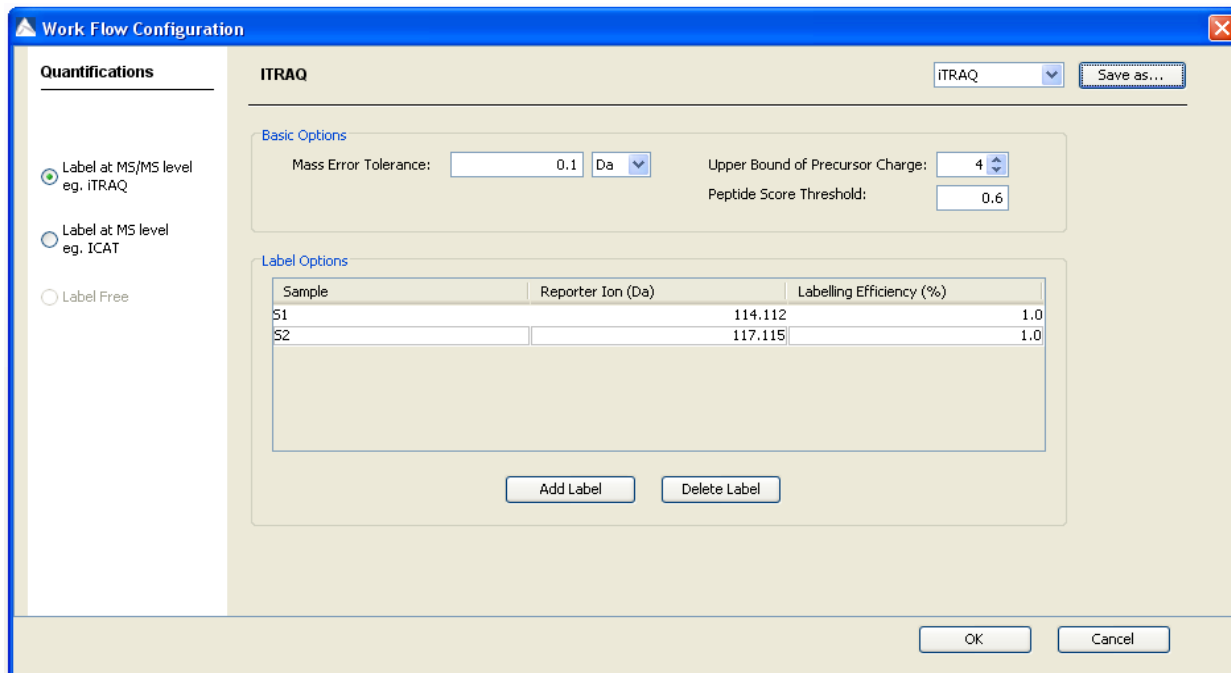
☐ Perform on-the-fly preprocessing (peak centroiding, charge deconvolution, and deisotope)

OK Cancel

Enter the PTMs as shown below by selecting “Set PTM”:



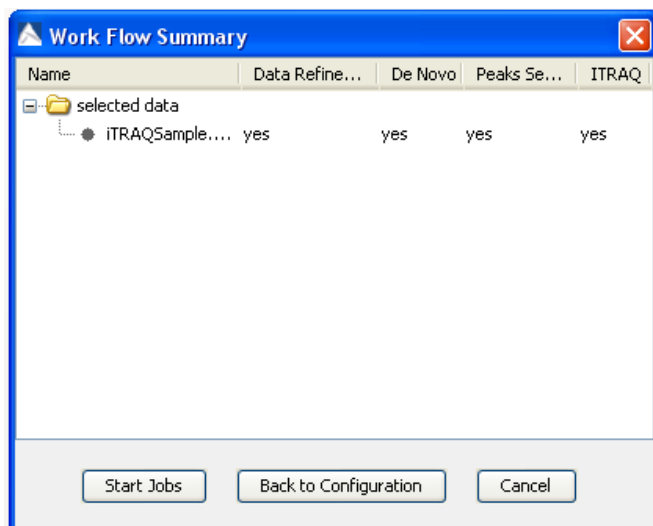
Check the 5th box and then click the “Quantification” button. Select “Label at the MS/MS level” from the left hand panel and enter the following parameters:



Note: To add information into the “Label Options” panel, click the “Add Label” button to add a new sample.

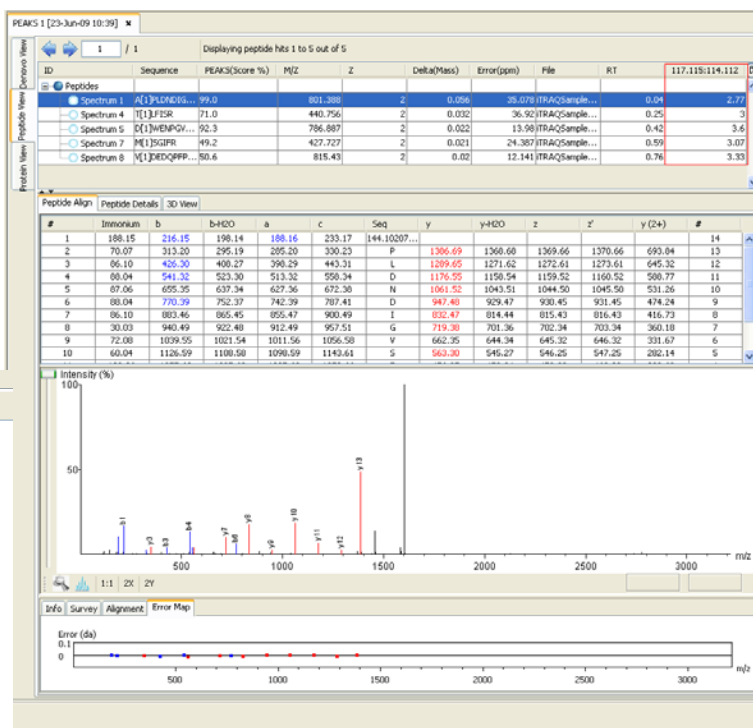
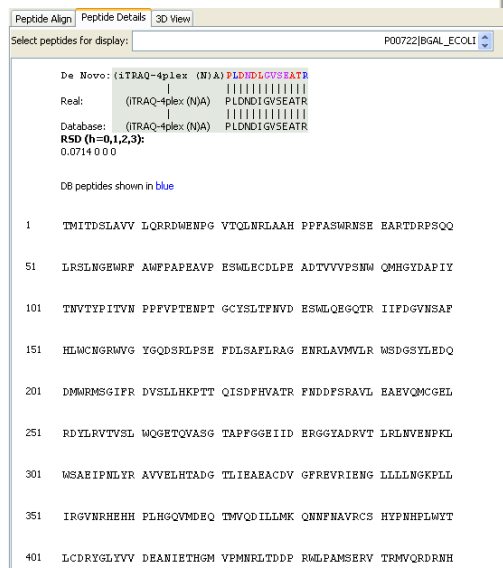
Click “OK” and then click “Start” in the workflow configuration panel. The following window will open:

Click “Start Jobs”.



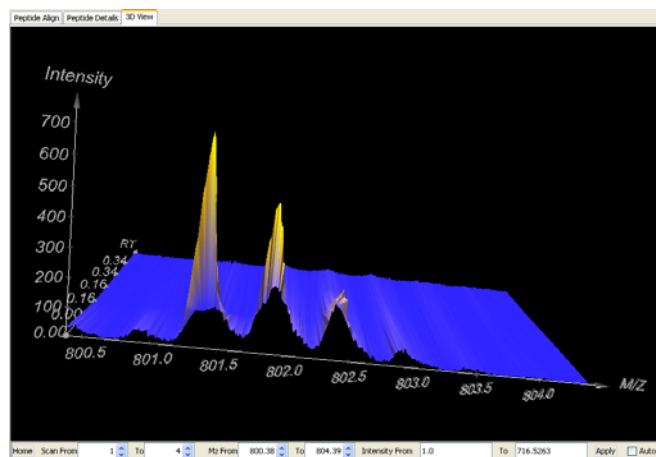
16.3.4 ITRAQ Results

Once completed, the protein quantification result will be displayed in the PEAKS Protein ID node . Double click on this node and the “Peptide View” tab will appear by default. The quantification results are listed as a “Ratio of 117.115:114.112”:



Click on the “Peptide Details” tab to see a simple alignment between the original *de novo* sequence, the peptide found in the database and the reconstructed sequence. At the bottom of the “Peptide Details” panel you will see where the selected peptide matches the protein highlighted in blue.

PEAKS provides the 3D View of each peptide feature for visual validation. Click on the 3D View tab. The panel along the bottom allows you to narrow in on the peptides that you would like to examine. You can specify a particular scan number range, m/z range or intensity range. Click the “Apply” button to change the 3D View to your specified values.



Select the “Protein View” tab. The quantification results are listed as a “Ratio of 117.115:114.112” and as “Standard Deviation of 117.115:114.112”. They are highlighted in the red box below. For example the relative protein ratio for the top ranked protein (Beta-galactosidase) is 3.01 with a standard derivation of 0.45:

PEAKS 1 [23-Jun-09 10:39]									
Accession	Display	PEAKS(Score %)	Coverage(%)	Query matched	Description	Ratio 117.115:114.112		SD 117.115:114.112	
P00722 BGAL_ECOLI		96.5	4.4	5	Beta-galactosidase...	3.01		0.45	
P00722 BGAL_...		96.5	4.39	5	Beta-galactosidase...	3.01		0.45	


Peptides									
ID	Sequence	PEAKS(Score %)	M/Z	Z	Mr(Calc)	Error(ppm)	RT	117.115:114.112	
Spectrum 1	A[1]PLDNDIGVSEATR	99.0	801.388	2	1600.8176	35.0785	0.04	2.77	
Spectrum 4	T[1]LFIISR	71.0	440.756	2	879.52997	36.9196	0.25	3.0	
Spectrum 5	D[1]WENPGVTQLNR	92.3	786.887	2	1571.7815	13.9797	0.42	3.6	
Spectrum 7	M[1]SGIFR	49.2	427.727	2	853.46027	24.3872	0.59	3.07	
Spectrum 8	V[1]DEDQPFPAVPK[2]	50.6	815.43	2	1628.8652	12.1407	0.76	3.33	

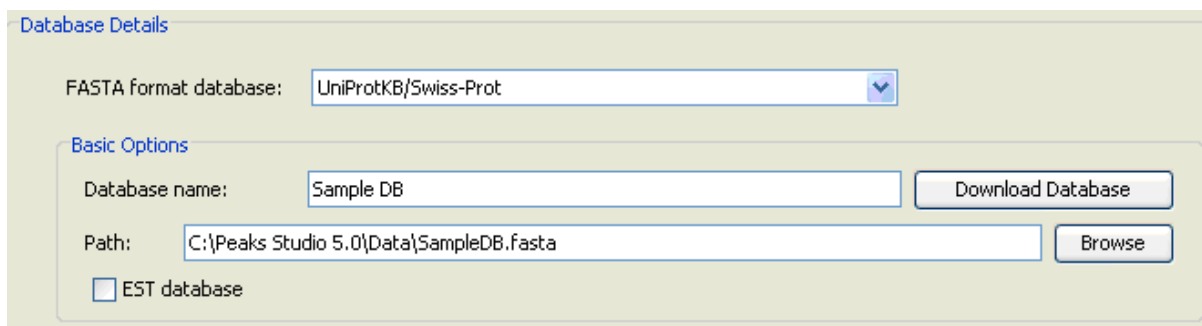
16.4 SILAC Walkthrough

Stable Isotope Labeling with Amino Acids in Cell Culture (SILAC) is a method to metabolically label proteins for relative quantitative comparison. One cell population is fed amino acids of normal isotopic composition; the other cell population is fed amino acids labeled with heavier isotopes. The heavy amino acids are incorporated into newly synthesized proteins, eventually completely replacing the cells’ proteins, such that labeling efficiency is near 100%. The cell populations are then mixed together and digested for MS analysis to determine differential protein abundances.

16.4.1 Setting up a Database

Before running the walkthrough data, you need to set up a database. So that this can be a quick process, we have provided you with a sample fasta database called “SampleDB.fasta” in your PEAKS program folder (C:\PEAKS Studio 5.1\Data).

Click on the configuration toolbar icon  or select “Configuration” from the “Tools” menu. Select “Database” from the left hand side of the window. Under “Database Details” enter the following information:

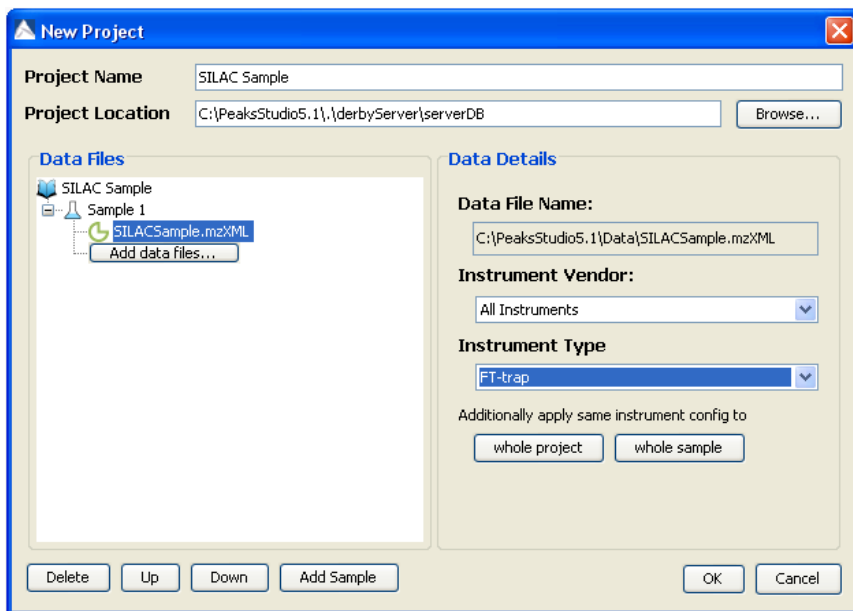


You do not need to change any of the other information listed. Click the “Add/Update” button and then click “OK”.

16.4.2 Creating a project


Click on the “Create new project” icon  or select “New project” from the “File” menu. Give a name to your project, for example “SILAC Sample”. Click Add data files. Locate “SILACSample.mzxml” from C:\PeaksStudio5.1\Data.

With the file highlighted, select, FT-trap from the “Instrument Type” drop-down list.

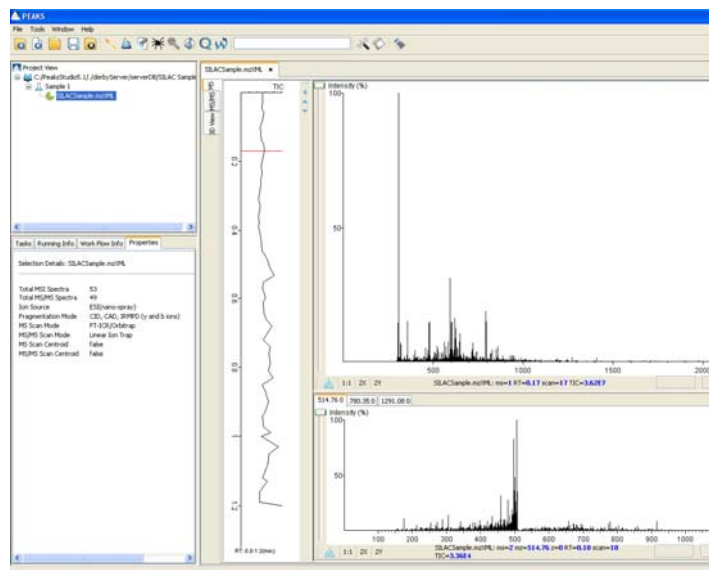


Click “OK”.

Double click on the SILAC sample file. Your “Main Processing Screen” should look something like this:

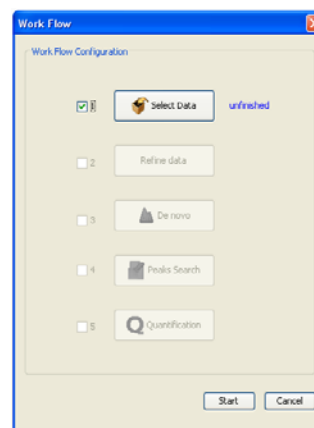
For ease of setup, this walkthrough will use the workflow tool in order to set all of the parameters at once. You will need to wait until the file has loaded (i.e. the file node looks like this: .

See section [16.2 3D View](#) to enable 3D view.

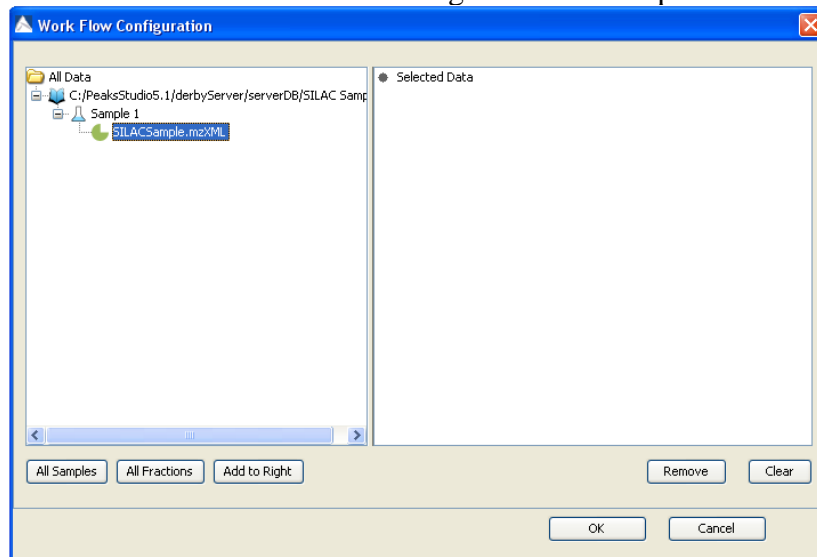


16.4.3 Setting up the Workflow Parameters

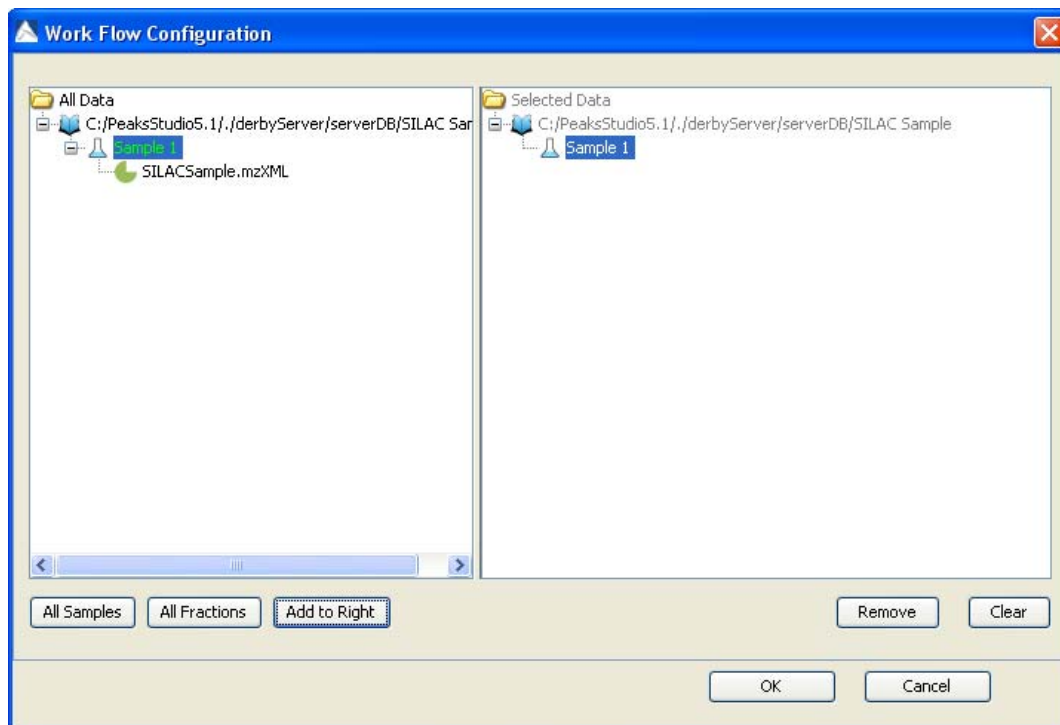
Click on the workflow button “”. The following window will appear:



Click “Select Data”. The following window will open:

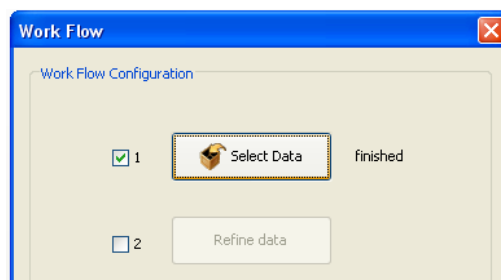


Select the SILACSample.mzXML file and then click “Add to Right”. The samples should now appear in the “Selected Data” panel as shown below:



Click “OK”.

The workflow window should now display “finished” beside the “Selected Data” button.



Check the 2nd box and then click the “Refine data” button. Enter the following parameters:

Work Flow Configuration

MS/MS Data Refinement

☐ Merge Scans

Retention time window (for raw files only): min

Precursor m/z error tolerance: Da ☐ ppm

☒ Correct Precursor Charge States

Min charge: Max charge:

☐ Filter Scans

Only keep scans satisfying:

☐ Precursor mass between and Da

☐ Retention time between and min

☐ Quality value greater than (suggest 0.65)

Data Preprocess (peak centroiding, charge deconvolution, and deisotope)

☐ no, already done ☒ yes ☐ no

OK Cancel

Click “OK”.

Check the 3rd box and then click the “*de novo*” button. Enter the following parameters:

Enter the PTMs as shown below by selecting “Set PTM”:

PTM Options

All PTM

Name	Mono mass	Residue site
Iodoacetic acid derivative	58.0055	[C]
Carboxymethyl	58.0055	[C]
K6	6.0201	[K]
Lipoyl	188.0230	[K]
Methyl ester	14.0156	[DC], [D]PC
Methylation	14.0156	[CRHCDKQ], [D]BN
O18 label	2.0042	[STY], [D]C
Propionamide	71.0371	[C]
Trimethylation	42.0470	[OURHCEVQ], [PAPVLM]...
Myristoylation	210.1964	[KC], [G]BN
N-acetyl diglycine cysteine	780.7250	[C]
N-isopropylcarbamate...	99.0684	[C]
N-succinyl-3-morphol...	127.0633	[D], [D]BN
Oxidation M	15.9949	[H]
Oxidation HMF	15.9949	[HMF]
Palmitylation	238.2297	[CSTK]
Phosphoserine/threonine	340.0858	[S]

☐ Show unlisted

Selected Fixed PTM

- Carbamidomethylation

Selected Variable PTM

- Oxidation M
- K6

Deselect Clear All Switch Type

OK Cancel

Click “OK”.

Work Flow Configuration

De Novo

Predefined parameters: SILAC

Error Tolerance

Parent ion: 0.1 Da Fragment ion: 0.8 Da

Enzyme

Trypsin

PTM

☒ Oxidation M

☒ K6

☐ Carbamidomethylation

Maximum allowed variable PTM per peptide: 0

General Options

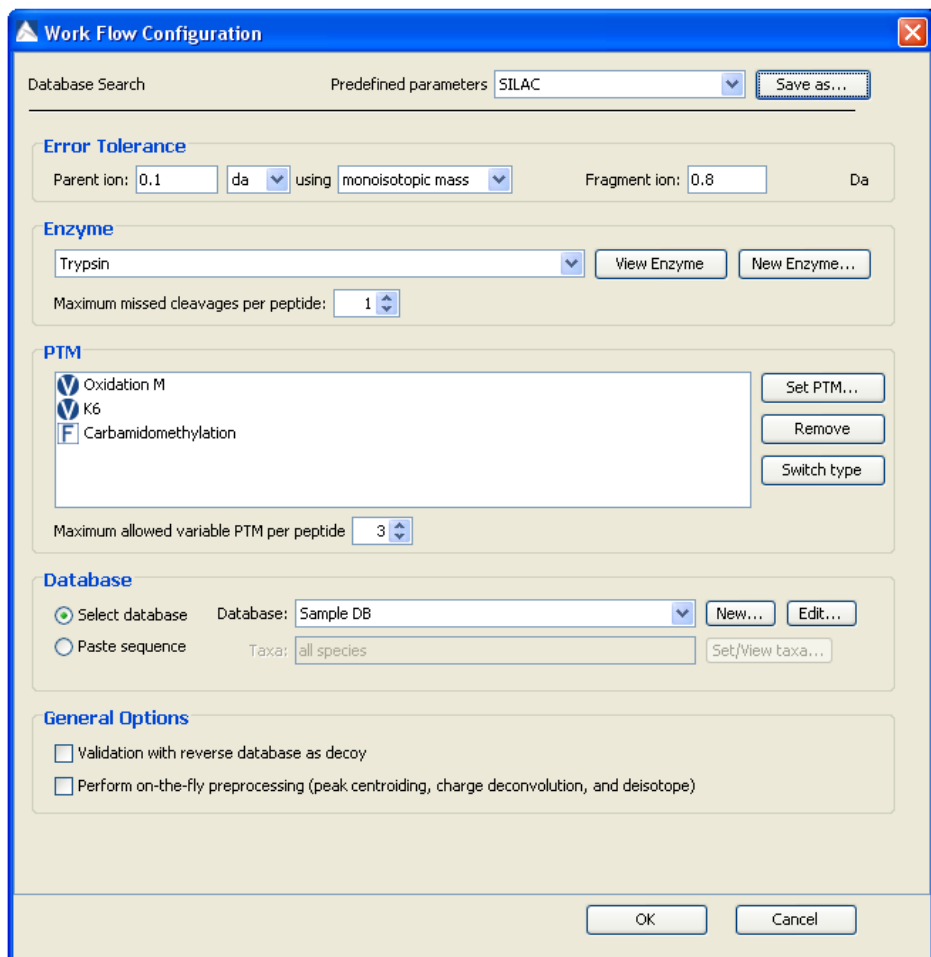
☒ Preprocess this data on-the-fly (peak centroiding, charge deconvolution and deisotope)

Report up to: 5 candidates per spectrum

OK Cancel

Check the 4th box and then click the “PEAKS Search” button. Enter the following parameters:

Enter the PTMs as shown below by selecting “Set PTM”:



Work Flow Configuration

Database Search: Predefined parameters: SILAC [Save as...]

Error Tolerance
 Parent ion: 0.1 da using monoisotopic mass Fragment ion: 0.8 Da

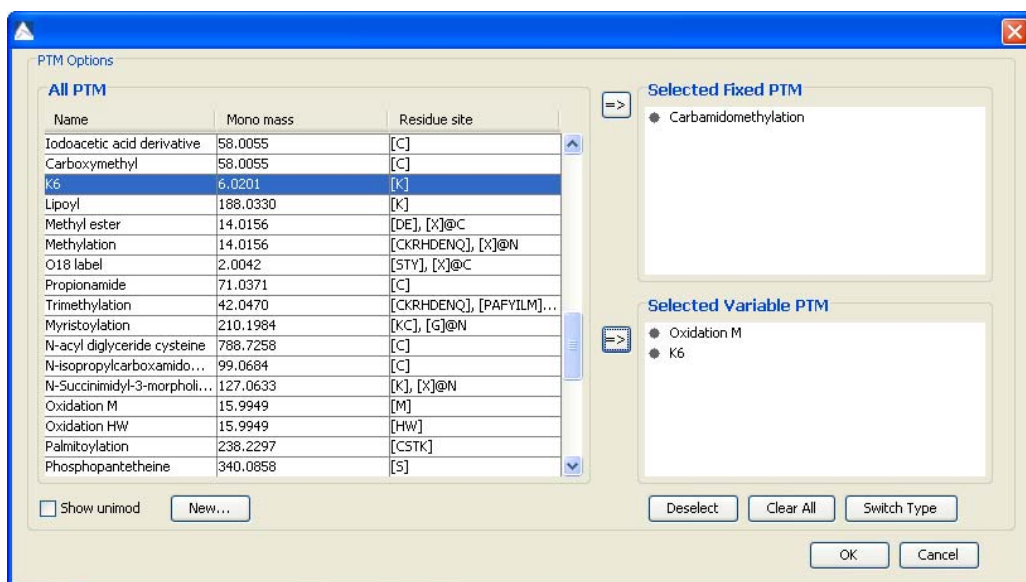
Enzyme
 Trypsin [View Enzyme] [New Enzyme...]
 Maximum missed cleavages per peptide: 1

PTM
☒ Oxidation M
☒ K6
☐ Carbamidomethylation
 [Set PTM...]
 [Remove]
 [Switch type]
 Maximum allowed variable PTM per peptide: 3

Database
☒ Select database Database: Sample DB [New...] [Edit...]
☐ Paste sequence Taxa: all species [Set/View taxa...]

General Options
☐ Validation with reverse database as decoy
☐ Perform on-the-fly preprocessing (peak centroiding, charge deconvolution, and deisotope)

[OK] [Cancel]



PTM Options

All PTM

Name	Mono mass	Residue site
Iodoacetic acid derivative	58.0055	[C]
Carboxymethyl	58.0055	[C]
K6	6.0201	[K]
Lipoyl	188.0330	[K]
Methyl ester	14.0156	[DE], [X]@C
Methylation	14.0156	[CKRHDEHQ], [X]@N
O18 label	2.0042	[STY], [X]@C
Propionamide	71.0371	[C]
Trimethylation	42.0470	[CKRHDEHQ], [PAFYILM]...
Myristoylation	210.1984	[KC], [G]@N
N-acyl diglyceride cysteine	788.7258	[C]
N-isopropylcarboxamido...	99.0684	[C]
N-Succinimidyl-3-morpholi...	127.0633	[K], [X]@N
Oxidation M	15.9949	[M]
Oxidation HW	15.9949	[HW]
Palmitoylation	238.2297	[CSTK]
Phosphopantetheine	340.0858	[S]

☐ Show unimod [New...]

Selected Fixed PTM
☒ Carbamidomethylation

Selected Variable PTM
☒ Oxidation M
☒ K6

[Deselect] [Clear All] [Switch Type]
 [OK] [Cancel]

Check the 5th box and then click the “Quantification” button. Select “Label at the MS level” from the left hand panel and enter the following parameters:

The **Work Flow Configuration** dialog box is shown with the **ICat** method selected. The **Quantifications** panel on the left has the option **Label at MS level eg. ICAT** selected. The **Basic Options** section contains the following settings:

- Mass Error Tolerance: 0.1 Da
- Retention Time Range: 1.0 min.
- Upper Bound of Precursor Charge: 4
- Peptide Score Threshold: 0.6

The **Label Options** section contains a table with the following data:

Sample	Added Mass	Residues	Labelling Efficiency (%)
Light		0.0K	1.0
Heavy		6.0K	1.0

Buttons for **Add Label** and **Delete Label** are located below the table. The **OK** and **Cancel** buttons are at the bottom right of the dialog.

Note: To add information into the “Label Options” panel, click the “Add Label” button to add a new sample.

Click “OK” and then click “Start” in the workflow configuration panel. The following window will open:

Click “Start Jobs”.


The **Work Flow Summary** dialog box displays the configuration for the workflow. It has a table with the following columns: **Name**, **Data Refinement**, **De Novo**, **Peaks Search**, and **ICAT**.

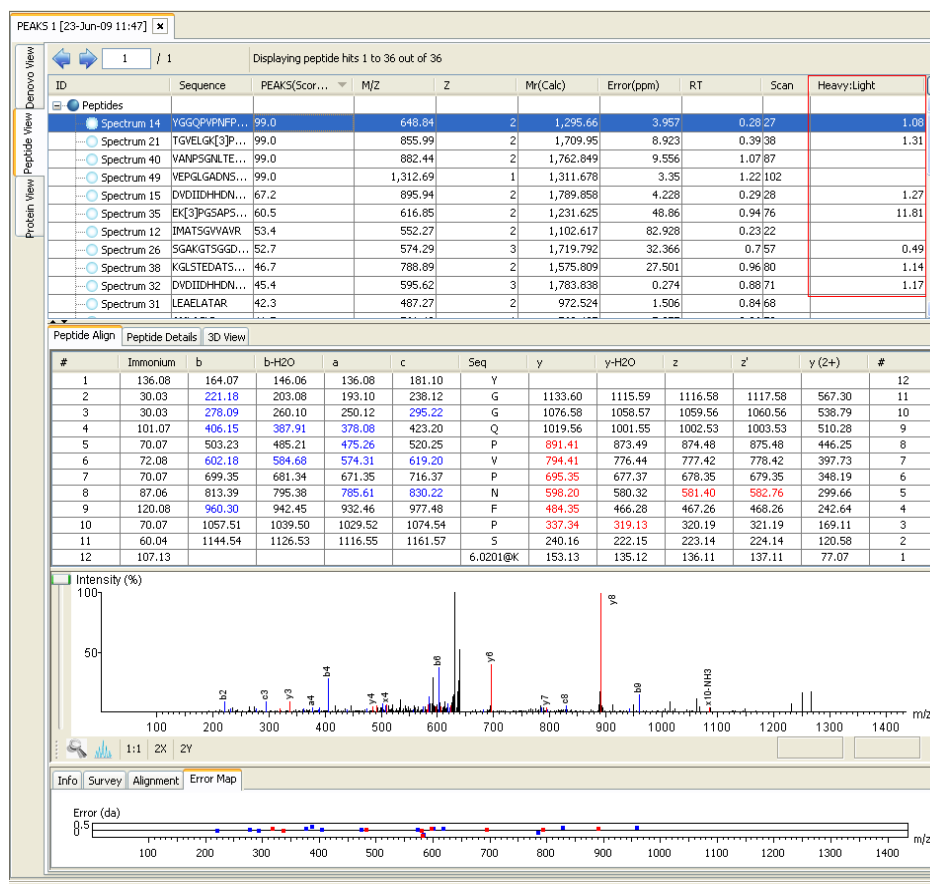
Name	Data Refinement	De Novo	Peaks Search	ICAT
selected data				
Sample 1	yes	yes	yes	yes

Buttons for **Start Jobs**, **Back to Configuration**, and **Cancel** are located at the bottom of the dialog.

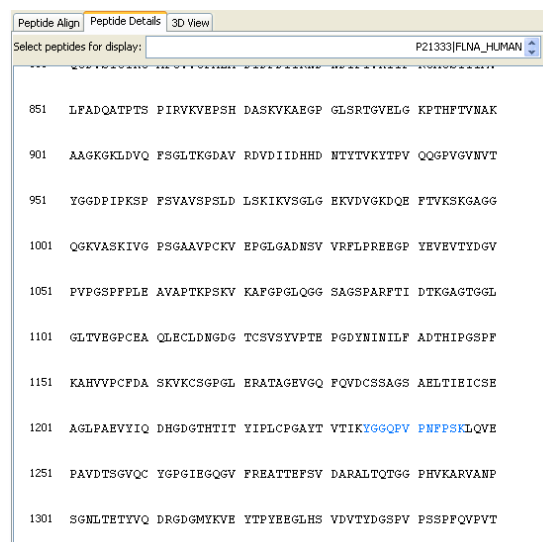
16.4.4 SILAC

Results

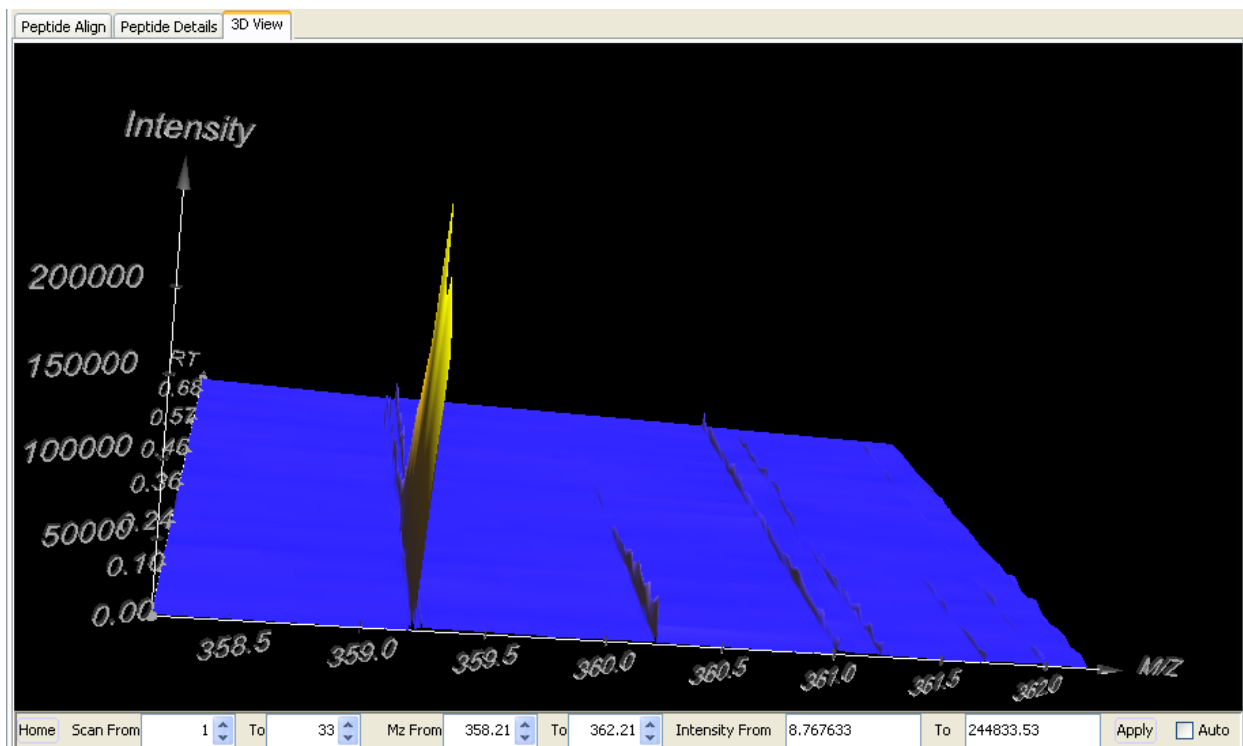
Once completed, the protein quantification result will be displayed in the PEAKS Protein ID node “”. Double click on this node the “Peptide View” results will appear by default:



Click on the “Peptide Details” tab to see where the selected peptide matches the protein highlighted in blue:



PEAKS provides the 3D View of each peptide feature for visual validation. Select the “3D View” tab. The panel along the bottom allows you to narrow in on the peptides that you would like to examine. You can specify a particular scan number range, m/z range or intensity range. Click the “Apply” button to change the 3D View to your specified values.



Select the “Protein View” tab. The SILAC quantification results are listed as a “Ratio Heavy: Light” and “Standard Deviation Heavy: Light”.







Accession	ID	Mass	Display	PEAKS(Score %)	Coverage(%)	Quer...	Description	Marked	Ratio Heavy:Light	SD Heavy:Light
<input checked="" type="radio"/> DB Search								<input type="checkbox"/>		
<input checked="" type="radio"/> P21333 FLN...	1	280736...		98.9	2.53	6	Filamin-A (Alph...	<input type="checkbox"/>	1.21	0.13
<input type="radio"/> Q72460 CLA...	3	169449...		19.2	0.78	2	CLIP-associatin...	<input type="checkbox"/>		
<input type="radio"/> P13942 COB...	4	171773...		18.8	1.15	2	Collagen alpha-...	<input type="checkbox"/>	-1	0
<input type="radio"/> Q15059 BR...	2	79541.586		24.3	1.65	1	Bromodomain-c...	<input type="checkbox"/>	11.81	0
<input type="radio"/> Q8N6N7 AC...	10	9790.278		15.2	15.91	1	Acyl-CoA-bindi...	<input type="checkbox"/>	-1	0
<input type="radio"/> P29400 CO4...	11	161043...		12.7	0.71	1	Collagen alpha-...	<input type="checkbox"/>	-1	0
<input type="radio"/> Q7Z2E3 APT...	12	40740.040		11.7	2.53	1	Aprataxin (For...	<input type="checkbox"/>	-1	0
<input type="radio"/> Q96EP1 CH...	13	73386.586		12.3	0.9	1	E3 ubiquitin-pr...	<input type="checkbox"/>		

Note that the top protein result is Human Filamin A, with a score of 98.89%. The ratio of Heavy: Light is highlighted for each protein in the red box below. For example, the highest ranked protein, Human Filamin-A has a ratio of 1.2 and a standard deviation of 0.13.

A ratio of -1 means that while a reporter ion was found for the selected peptide, a ratio could not be calculated for the protein as the peptide did not meet the required score criteria (in this case we specified that the peptide score had to be greater than 0.6). A blank box in the ratio column means that a reporter ion for the selected peptide was not found.

Accession	ID	Mass	Display	PEAKS(Score %)	Coverage(%)	Ratio Heavy:Light	SD Heavy:Light	Query matched	Marked	Description
DB Search										
Q28333 FILN_HU	61	28025.52		98.95	2.53	1.18	0.1	6		Filamin-A (Alpha-Filamin)
P04264 K2C1_HU	62	66017.45		34.7	4.97			3		Keratin, type II cytoske...
Q9Y283 INVS_HU	64	117826.305		31.33	0.94	19.83	0	1		Inversin (Inversion of e...
Q8M6N7 ACR07_J	66	9790.278		15.6	15.91	1.09	0	1		Acyl-CoA-binding domai...
Q96EU6 CF153_J	67	29823.105		13.72	4.63	10.49	0	1		UPF0399 protein C6orf...
P13497 BMP1_HU	68	111248.77		11.87	0.81	0.49	0	1		Bone morphogenetic pr...
Q72460 CLAP1_H	69	169449.78		10.32	0.65			1		CLIP-associating protei...
Q8MS20 AADAT_J	70	47351.652		9.85	3.29			1		Kynurenine/alpha-amin...
Q96EP1 CHFR_HU	71	73386.586		8.33	0.9			1		E3 ubiquitin-protein liga...

There are four peptide features are identified and calculated for quantification of Human Filamin-A. Two other peptides (Spectrum 40 and 49) were used to identify the protein but were not identified as features.

Peptides	Coverage	Tool Box							
ID	Sequence	PEAKS(Score %)	M/Z	Mr(Calc)	Error(ppm)	RT	Scan	Heavy:Light	
 Spectrum 14	YGGQVPVPNFPSK[3]	99.0	648.84	1295.6604	3.957	0.28	27	1.08	
 Spectrum 15	DVDIIDHHDNTYTVK[3]	67.3	895.94	1789.8579	4.2285	0.29	28	1.27	
 Spectrum 21	TGVELGK[3]PTHFTVNAK[3]	99.0	855.99	1709.9502	8.9234	0.39	38	1.31	
 Spectrum 32	DVDIIDHHDNTYTVK	45.4	595.62	1783.8376	0.2737	0.88	71	1.17	
 Spectrum 40	VANPSGNLTETYVQDR	99.0	882.44	1762.8486	9.5559	1.07	87		
 Spectrum 49	VEPGLGADNSVVR	99.0	1312.6899	1311.6782	3.3503	1.22	102		


16.5 Label Free Quantification

Label Free quantification relies on the changes in analyte signals directly reflecting their concentrations in one sample relative to another. This technology employs overall spectral intensity normalization by interpreting signals of molecules that do not change concentration from sample to sample. By comparing two or more spectra, PEAKS can determine the constant intensity ratio between the unchanging analytes forms the basis for identifying the non-changing concentrations, making spiking unnecessary.

Below are instructions to take you through a label free dataset. Please be aware that the data files are fairly large and thus the computing time will be increased compared to the other walkthrough datasets explained in this manual.

16.5.1 Setting up a Database

Before running the walkthrough data, you need to set up a database. So that this can be a quick process, we have provided you with a sample fasta database called “SampleDB.fasta” in your PEAKS program folder (C:\PEAKS Studio 5.1\Data).

Click on the configuration toolbar icon  or select “Configuration” from the “Tools” menu. Select “Database” from the left hand side of the window. Under “Database Details” enter the following information:

Database Details

FASTA format database: UniProtKB/Swiss-Prot

Basic Options


Database name: Sample DB Download Database

Path: C:\Peaks Studio 5.0\Data\SampleDB.fasta Browse

☐ EST database

You do not need to change any of the other information listed. Click the “Add/Update” button and then click “OK”.

16.5.2 Creating a Project

Click on the “Create new project” icon  or select “New project” from the “File” menu. Give a name to your project, for example “Label free walkthrough”. Click Add data files. Locate LFQSample1.mzXML. Then click “Add Sample” and Sample 2 will appear. Click “Add data files” to add “LFQSample2.mzXML” to Sample 2.

Note: For label free quantification to function in PEAKS, you need to have at least 1 sample with at least 1 file/fraction in each sample.

With one of the raw files highlighted, select, FT-trap from the “Instrument Type” drop-down. Then click “whole project” to apply these parameters to the rest of the project.

New Project

Project Name: Label free walkthrough

Project Location: C:\PeaksStudio5.1\.\derbyServer\serverDB Browse...

Data Files

- Label free walkthrough
 - Sample 1
 - LFQSample1.mzXML
 - Add data files...
 - Sample 2
 - LFQSample2.mzXML
 - Add data files...

Data Details

Data File Name: :s and Settings\user\Desktop\LFQSample1.mzXML


Instrument Vendor: All Instruments

Instrument Type: FT-trap

Additionally apply same instrument config to: whole project whole sample


Delete Up Down Add Sample OK Cancel

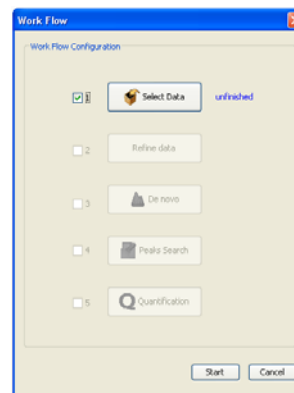
Click “Ok”.

For ease of setup, this walkthrough will use the workflow tool in order to set all of the parameters at once. You will need to wait until the files have loaded (i.e. the file node looks like this: ).

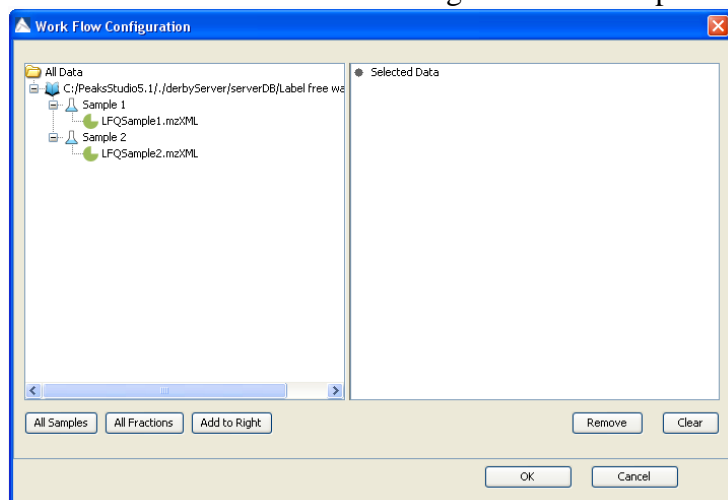
See section [16.2 3D View](#) to enable 3D view.

16.5.3 Setting up the Workflow Parameters

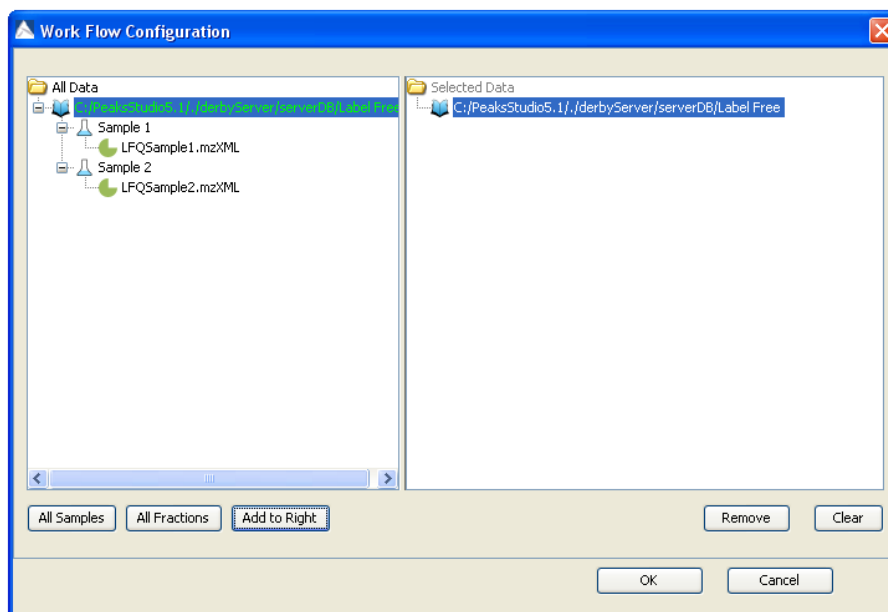
Click on the workflow button “”. The following window will appear:



Click “Select Data”. The following window will open:



Click “All Samples” and then “Add to Right”. The samples should now appear in the “Selected Data” panel as shown:



Click “OK”.

The workflow window should now display “finished” beside the “Selected Data” button.

Check the 2nd box and then click the “Refine data” button. Enter the following parameters:

Click “OK”.

Work Flow Configuration

MS/MS Data Refinement

☐ Merge Scans
Retention time window (for raw files only): min
Precursor m/z error tolerance: ☐ Da ☐ ppm

☒ Correct Precursor Charge States
Min charge: 1 Max charge: 4

☒ Filter Scans
Only keep scans satisfying:
☐ Precursor mass between and Da
☐ Retention time between and min
☒ Quality value greater than 0.65 (suggest 0.65)

Data Preprocess (peak centroiding, charge deconvolution, and deisotope)
☐ no, already done ☒ yes ☐ no

OK Cancel

Work Flow Configuration

De Novo

Predefined parameters: label free Save as...

Error Tolerance
Parent ion: 0.1 Da Fragment ion: 0.8 Da

Enzyme
 Semi Trypsin View Enzyme New Enzyme...

PTM
☒ Oxidation M ☒ Phosphorylation-STY ☐ Carbamidomethylation
 Set PTM... Remove Switch type

Maximum allowed variable PTM per peptide: 3

General Options
☐ Preprocess this data on-the-fly (peak centroiding, charge deconvolution and deisotope)
Report up to 5 candidates per spectrum

OK Cancel

Check the 3rd box and then click the “De novo” button. Enter the following parameters:

Enter the PTMs as shown below by selecting “Set PTM”:

Click “OK”.

PTM Options

Name	Mono mass	Residue site
Trimethylation	42.0470	[CKRHDENQ], [PAFYILM]...
Myristoylation	210.1984	[KC], [G]@N
N-acyl diglyceride cysteine	788.7258	[C]
N-isopropylcarboxamido...	99.0684	[C]
N-Succinimidyl-3-morphol...	127.0633	[K], [X]@N
Oxidation M	15.9949	[M]
Oxidation HW	15.9949	[HW]
Palmitoylation	238.2297	[CSTK]
Phosphopantetheine	340.0858	[S]
Phosphorylation	79.9663	[STYHCDR]
Phosphorylation-STY	79.9663	[STY]
Pyridoxal phosphate	229.0140	[K]
Pyro-glu from E	-18.0106	[E]@N
Pyro-glu from Q	-17.0265	[Q]@N
Pyrrolidone carboxylic acid	-17.0266	[Q]@N
S-carbamoylmethylcystei...	-17.0265	[C]@N
Sodium adduct	21.9819	[DE], [X]@C

☐ Show unimod New...

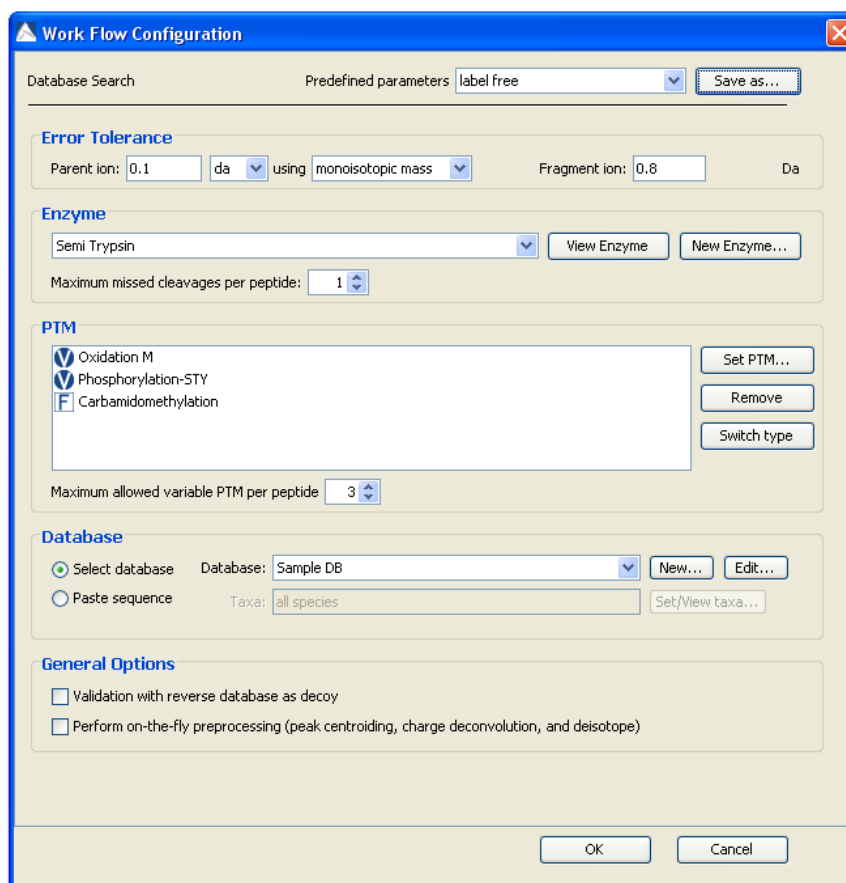
Selected Fixed PTM
☒ Carbamidomethylation

Selected Variable PTM
☒ Oxidation M
☒ Phosphorylation-STY

Deselect Clear All Switch Type

OK Cancel

Check the 4th box and then click the “PEAKS Search” button. Enter the following parameters:



Work Flow Configuration

Database Search: Predefined parameters: label free [Save as...]

Error Tolerance
 Parent ion: 0.1 da using monoisotopic mass Fragment ion: 0.8 Da

Enzyme
 Semi Trypsin [View Enzyme] [New Enzyme...]
 Maximum missed cleavages per peptide: 1

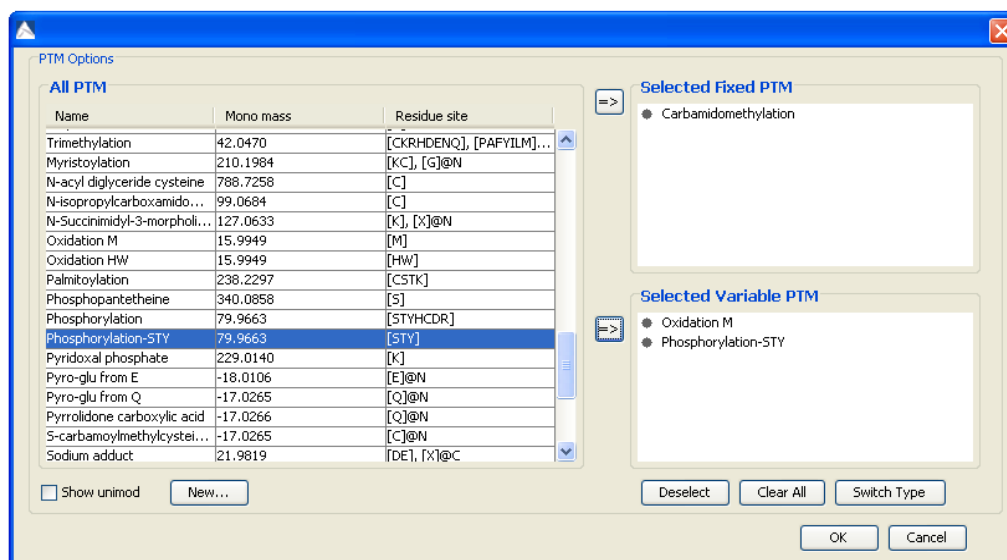
PTM
☒ Oxidation M
☒ Phosphorylation-STY
☐ Carbamidomethylation
 [Set PTM...]
 [Remove]
 [Switch type]
 Maximum allowed variable PTM per peptide: 3

Database
☒ Select database Database: Sample DB [New...] [Edit...]
☐ Paste sequence Taxa: all species [Set/View taxa...]

General Options
☐ Validation with reverse database as decoy
☐ Perform on-the-fly preprocessing (peak centroiding, charge deconvolution, and deisotope)

[OK] [Cancel]

Enter the PTMs as shown below by selecting “Set PTM”:



PTM Options

Name	Mono mass	Residue site
Trimethylation	42.0470	[CKRHDEHQ], [PAFYILM]...
Myristoylation	210.1984	[K], [G]@N
N-acyl diglyceride cysteine	788.7258	[C]
N-isopropylcarboxamido...	99.0684	[C]
N-Succinimidyl-3-morphol...	127.0633	[K], [X]@N
Oxidation M	15.9949	[M]
Oxidation HW	15.9949	[HW]
Palmitoylation	238.2297	[CSTK]
Phosphopantetheine	340.0858	[S]
Phosphorylation	79.9663	[STYHCDR]
Phosphorylation-STY	79.9663	[STY]
Pyridoxal phosphate	229.0140	[K]
Pyro-glu from E	-18.0106	[E]@N
Pyro-glu from Q	-17.0265	[Q]@N
Pyrrolidone carboxylic acid	-17.0266	[Q]@N
S-carbamoylmethylcyste...	-17.0265	[C]@N
Sodium adduct	21.9819	[DE], [X]@C

☐ Show unimod [New...]

Selected Fixed PTM
☒ Carbamidomethylation

Selected Variable PTM
☒ Oxidation M
☒ Phosphorylation-STY

[Deselect] [Clear All] [Switch Type]
 [OK] [Cancel]

Check the 5th box and then click the “Quantification” button. Select “Label Free” from the left hand panel and enter the following parameters:

Work Flow Configuration

Quantifications

- ☐ Label at MS/MS level eg. iTRAQ
- ☐ Label at MS level eg. ICAT
- ☒ Label Free

Label free walkthrough

Label Free label free Save as...

Mass Error Tolerance: Da Peptide Score: Min. Max.

min. Protein Score: Min. Max.

Project Name	Sample Name	Fraction Number	File Name	Protein ID
Label free walkthrough				
	Sample 1			
		1	LFQSample1.mzXML	from work flow
	Sample 2			
		1	LFQSample2.mzXML	from work flow

OK Cancel

Click “OK” and then click “Start” in the workflow configuration panel. The following window will open:

Click “Start Jobs”.

Work Flow Summary

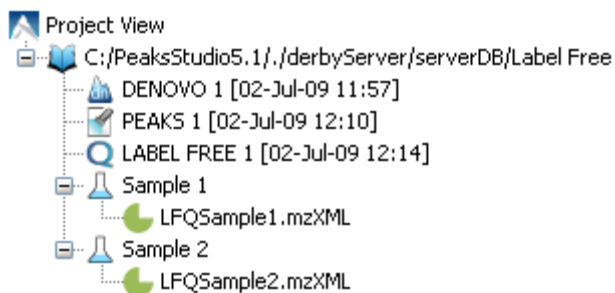
Name	Data Refinement	De Novo	Peaks Search	Label Free
selected data				
Label Free yes	yes	yes	yes	yes

Start Jobs Back to Configuration Cancel

16.5.4 Label Free Results

Once the workflow analysis is complete, you should see the following in your “Project View” panel:

Double click on the label free node. The identified proteins will appear in the top panel, with homologous proteins clustered together. The ratio of Sample 1: Sample 2 appears in the “Ratio” column and the standard deviation of Sample 1: Sample 2 appears in the “SD” column.



In the example below the standard deviation is 0 as there is only one supporting peptide for the protein shown.

Accession	ID	Mass	Score	Coverage	Matched	Description	Ratio	SD	Marked
P60709 ACTB_HUMAN	4	41,736.754	7.4	3.2		1 Actin, cytoplasmic 1 ...	1.00 : 1.05	0.00 : 0.00	<input type="checkbox"/>
P68133 ACT5_HUMAN	9	42,051.051	7.4	3.18		1 Actin, alpha skeletal ...	1.00 : 1.05	0.00 : 0.00	<input type="checkbox"/>
P63267 ACTH_HUMAN	6	41,876.898	7.4	3.19		1 Actin, gamma-enteri...	1.00 : 1.05	0.00 : 0.00	<input type="checkbox"/>
P63261 ACTG_HUMAN	5	41,792.859	7.4	3.2		1 Actin, cytoplasmic 2 ...	1.00 : 1.05	0.00 : 0.00	<input type="checkbox"/>
P68032 ACTC_HUMAN	8	42,018.992	7.4	3.18		1 Actin, alpha cardiac ...	1.00 : 1.05	0.00 : 0.00	<input type="checkbox"/>
P62736 ACTA_HUMAN	7	42,008.973	7.4	3.18		1 Actin, aortic smooth ...	1.00 : 1.05	0.00 : 0.00	<input type="checkbox"/>

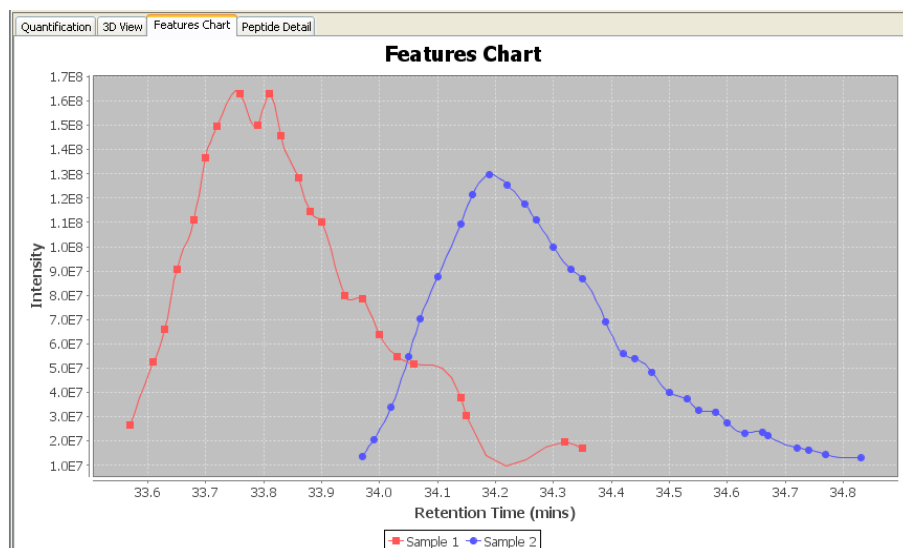
The supporting peptide is shown under the “Peptides” tab. The retention time is shown for the specific peptide as well as the peptide ratio from Sample 1: Sample 2.

Sequence	Peaks QScore	M/Z	Z	Retention Time	Ratio
Support Peptides					
DSYVGDEAQSK	7.4	599.765	2	[33.57, 34.35]	1.00 : 0.80
Outlier					
Multi Cluster Supports					

Click on the “+” beside the “Outlier” folder to see the peptides that were not included in the ratio.

To see which peptides were used to identify the protein during the PEAKS protein ID search, select the “Coverage” tab. The entire sequence of the protein is shown and the matching peptides are highlighted in blue. In this example the total matched part accounts for 9.392% of the protein. This information can be found in the “Coverage” column in the “Protein View” panel.

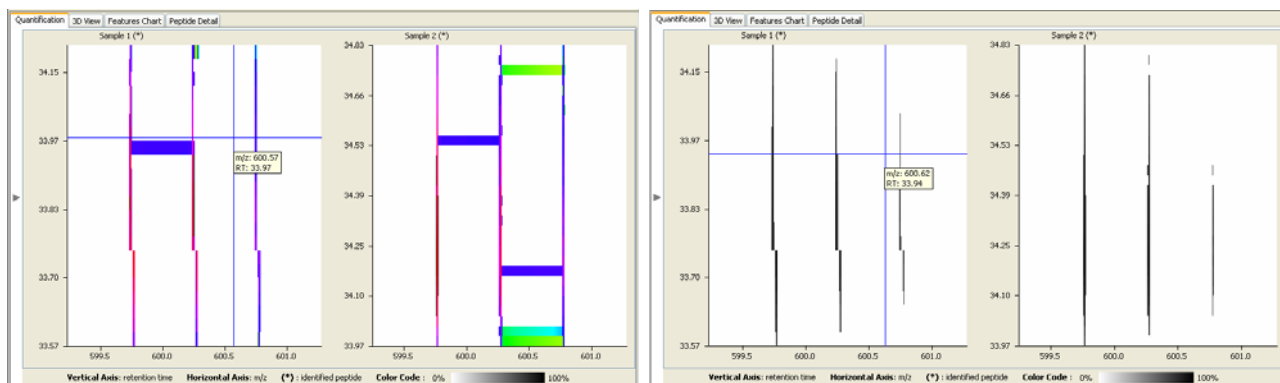
The features chart will appear by default in the bottom panel:



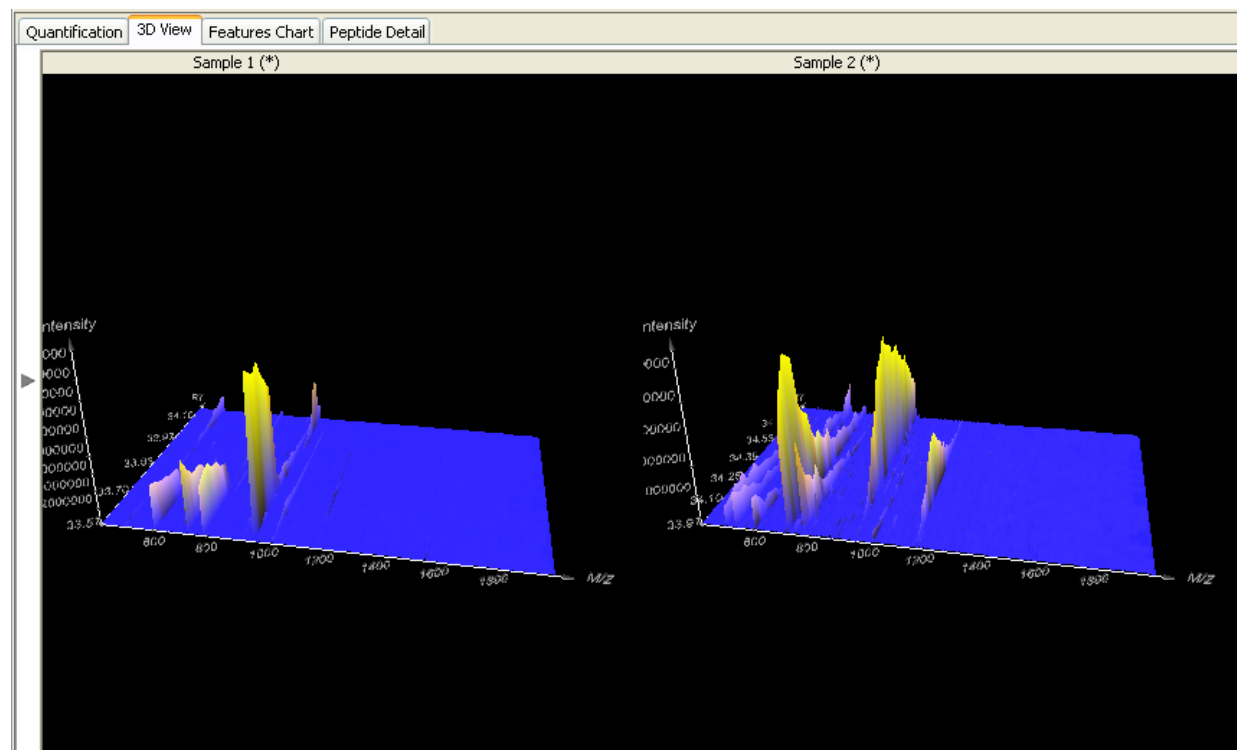
To see exactly where the selected supporting peptide corresponds to the protein sequence, select the “Peptide detail” tab. The entire sequence of the protein is shown and the selected supporting peptide is highlighted in blue.

Clicking on the “Quantification” tab will display a 2D view of the features. Move your cursor around the map to display the m/z ratio and retention time.

Right click your mouse and select “Show features” to display a 2D heat map. When viewing the heat map in color, red represents high intensity and yellow represents low intensity. The grayscale heat map displays high intensity in black while white represents low intensity.



Click on the “3D View” tab to display a 3D View of the peptide features for sample 1 and sample 2. Intensity is displayed on the y-axis, m/z on the x-axis and retention time on the z-axis. Click on the image and move the cursor to rotate.



17. References

de novo

Bin Ma, Kaizhong Zhang, Christopher Hendrie, Chengzhi Liang, Ming Li, Amanda Doherty-Kirby, Gilles Lajoie. PEAKS: Powerful Software for Peptide *de novo* Sequencing by MS/MS. Rapid Communications in Mass Spectrometry, 17(20):2337-2342. 2003.

SPIDER

Y. Han, B. Ma, and K. Zhang. SPIDER: Software for Protein Identification from Sequence Tags Containing *de novo* Sequencing Error. Journal of Bioinformatics and Computational Biology 3(3):697-716. 2005.

Quantification

Yang, W., Chen, W., Rogers, I., Ma, B., Bendall, S., Lajoie, G., Smith, D., PEAKS Q: Software for MS-based quantification of stable isotope labeled peptides. Bioinformatics Solutions Inc., Genome BC Proteomics Centre, University of Western Ontario. ASMS 2006 poster WP531.

18. Appendix

18.1 Terminology and Abbreviations Glossary

a-ions: an N-terminal fragment holding at least one charge. This is a fragment of the peptide derived from b-ions. The a-ion's mass will be the sum of the masses of the N terminal group, plus the intervening neutral amino acid residues, minus the mass of carbon monoxide.

b-ions: an N-terminal fragment holding at least one charge. This is a prefix fragment of the peptide. The b-ion's mass will be the sum of the masses of the N terminal group, plus the intervening neutral amino acid residues.

BSI (Bioinformatics Solutions Inc.): The makers of PEAKS and other fine bioinformatics software.

c-ions: an N-terminal fragment holding at least one charge. This is a prefix fragment of the peptide. The c-ion's mass will be the sum of the masses of the N terminal group, plus the intervening neutral amino acid residues, plus the mass of ammonia.

Deconvolution: rearrangement of the spectrum to show each monoisotopic peak as if it were singly charged. Thus, to reposition them on the scale, PEAKS multiplies the m/z of ion's that were doubly charged by two minus the mass of 1 H. Note that the deconvolved scale PEAKS shows is 'at +1.'

Fixed modification: selecting a post-translational modification as a fixed modification tells PEAKS that this modification is applied to all occurrences of the residue(s) on which the PTM can occur.

Enzyme: Biomolecules that catalyze chemical reactions, including the digestion of proteins.

ESI (Electrospray Ionization): A method for ionizing a sample into the mass spectrometer.

m/z : mass to charge ratio.

MALDI (Matrix-Assisted Laser Desorption/Ionization): A method for ionizing a sample into the mass spectrometer. This has a characteristic effect of producing singly charged ions.

Mass accuracy: this refers to the accuracy of data obtained from a given mass spectrometer. On a spectrum, how close the peaks are to the actual masses of the ions they represent.

PTM (Post-Translational Modification): A newly translated protein may differ from its final form as a result of processing by various enzymes in the cellular environment. This change is referred to as a post-translational modification. Since PTMs change the mass of residues, it must be accounted for when sequencing peptides by mass spectrometry.

Built-in PTM: PEAKS comes equipped with a library of possible post-translational modifications. These can be incorporated into a *de novo* analysis at the click of a button.

Customized PTM: If the post-translational modification you are looking for is not in the PEAKS PTM set, you may create our own entry, or modify an existing one. This will appear as a customized PTM in the set.

PTM library: A listing of all possible (built-in and custom entered) post-translational modifications that PEAKS can use as a part of its analysis.

Residue: as used in this manual, a residue refers to what remains of an amino acid once it has become part of a peptide, or peptide fragment. In this manual, residues are referred to by their original amino acid names.

Resolution: refers to the resolving power of an instrument. On a spectrum, this is reflected by how close together two peaks can be and still be resolved.

Variable modification: selecting a post-translational modification as a variable modification tells PEAKS that this modification may or may not be applied to the residue(s) on which the PTM can occur.

x-ions: a C terminal fragment holding at least one charge. The x-ion's mass will be the sum of the masses of the C terminal group, plus the intervening neutral amino acid residues, plus the mass of carbon monoxide.

y-ions: a C terminal fragment holding at least one charge. The y-ion's mass will be the sum of the masses of the C terminal group, plus the intervening neutral amino acid residues, plus the mass of 2 H.

z-ions: a C terminal fragment holding at least one charge. The z-ion's mass will be the sum of the masses of the N terminal group, plus the intervening neutral amino acid residues, after subtracting the mass of ammonia.

18.2 Toolbars


18.2.1 Main Window Toolbar





New Project: Clicking the New Project button will allow users to create a brand new project, offering organization to any study.





Add Data: This allows you to add additional data to any existing project.


 **Open Recent Project:** The quickest method to recalling an existing project. Users are instantly directed to a selection of recently created, modified or viewed PEAKS projects.


 **Close Project:** When a project is complete or a user temporarily no longer is working on a particular project, users can click this icon to remove it from their Project View.


 **Save As:** The project method of PEAKS allows for all results to be automatically constantly saved. However, if a user wishes to save a particular application under a new heading this is the most convenient and appropriate way to do so. Just click the icon and enter the new title in the field provided. The file will be saved in the ANZ format. Press this after selecting a data file in the Peptide Data Frame.


 **Print:** Whether a user desires to print their ms spectrum views to complement a publication or print matrices, the print feature offers a straight forward connection to any printer configured to the user's computer.


 **Export:** Easily export the spectrum view, ion table, or a picture (bmp, gif, or jpg format) with ions, masses, PEAKS and peptides marked.


 **Exit:** To exit from the PEAKS software safely, select this icon or press on the keyboard the 'Control' key simultaneously with the letter 'Q'.


 **Data Refine:** Merge scans of the same peptide, remove noise spectra, preprocess within each MS/MS spectrum and recover peptide charge state. The data refinement options dialogue will allow us to choose and to set parameters for each of these refinement tools.


 **de novo:** Perform auto *de novo* for a selected data file, spectrum or project. Press this after selecting one or more data files (or spectra) in the Peptide Data Frame. An auto *de novo* options dialogue will allow users to set parameters before beginning.


 **PEAKS Search:** Perform protein identification on a selected data project. Press this after selecting one or more data files (or spectra) in the Peptide Data Frame. A protein identification options dialogue will allow users to set parameters.


 **SPIDER Search:** Peptide homologue search tool.


 **PTM Finder:** A PTM finder search can be performed on any PEAKS Protein ID and allows the user to be able to identify more PTMs in less time.


 **inChorus Search:** inChorus, a meta protein identification tool, can be used to compare and validate results, calling upon such search engines as Mascot, OMSSA, PEAKS, Sequest and X!Tandem, as well as SPIDER. Click this icon to set a task to have inChorus perform a user's analysis easily.

 **Quantification:** Where users require abundance ratio results, PEAKS Q provides insight. Click here to analyze results from ICAT, iTRAQ, Label Free, SILAC, N-terminal and User Defined Labeling techniques.

 **Workflow:** A batch process, where users can analyze several samples at once by performing data refine, *de novo* sequencing, protein ID and quantification.







 **Configuration:** Set up enzymes, post translational modifications, databases, instruments and parameters.

 **Preferences:** Organize general properties such as directories, search engines and ion editing capabilities.


 **Mass Calculator:** The Mass Calculator is a simple tool to help us determine the molecular weight of a peptide. Clicking this icon will make the mass calculator appear.

18.2.2 Project View





The Project View allows optimal organization and greater control when managing multiple files at once. As different projects may be active at once, it is important to understand the different categories and levels presented by PEAKS.

-  Tree Root
-  Project Node
-  Filter Result
-  Sample Node
-  File Node (when file is loading)
-  File Node (when file is loaded)

Search Engines: Within the Project View, these icons are used to represent which method/search engine a particular file was interpreted by.


-  Mascot
-  OMSSA
-  Sequest
-  X!Tandem

18.2.3 Main Processing Window Toolbar

-  Profile Mode
-  Peak Mode
-  1:1 Return to original size
-  2X Zoom X-axis

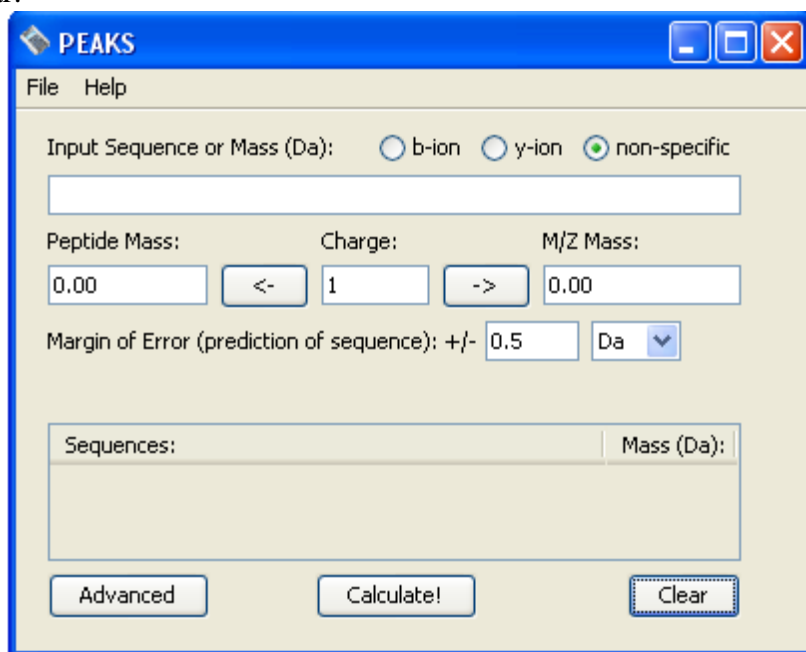
2Y Zoom Y-axis

18.3 Mass Calculator

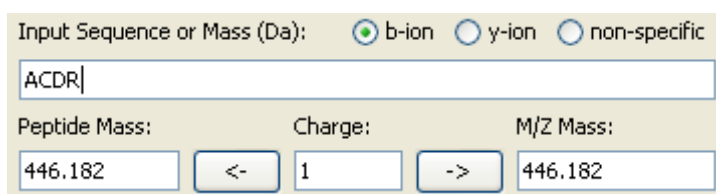
Click the Mass Calculator icon on the toolbar 
Or
choose Mass Calculator from the Tools menu.

The following window will appear:

In order to use the mass calculator input the peptide sequence or enter the mass in Daltons. Indicate if the sequence contains b- or -y ions or if your search is non-specific.



For example select “b-ion”, input the sequence ACDR and click the “Calculate” button. You should see the following:



You can change the charge and use the arrow on the right to calculate the precursor mass or use the arrow on the left to calculate the peptide mass.

Or for another example, change the margin of error to 1 Da, input a mass of 146 Da and click the “Calculate” button. In the sequences box you should see the following predicted sequences:

Sequences:	Mass (Da):
GS	144.0535
SG	144.0535

Use the “Clear” button to start again.

18.3.1 Advanced Options

Clicking on the “Advanced” button will display the following:

Advanced Options

A C D E F G H I K L

M N P Q R S T V W Y

Water Proton PTM Back Clear

Modification	Mass	Residues	Include
Acetylation (N-te...	42.010567	[ADCEQGILM...	
Acetylation (K)	42.010567	[K]@Anywhere	
Amidation	-0.984016	[ARNDCEQG...	
Applied Biosyste...	442.225	[C]@Anywhere	
Applied Biosyste...	450.2752	[C]@Anywhere	
Applied Biosyste...	227.12698	[C]@Anywhere	

☐ unimod Delete PTM New PTM Edit PTM

You will now be able to add various modifications to your amino acid sequence. For example, enter A, M, then select “Oxidation on M” from the PTM list and click the “PTM” button. Then click the “Calculate” button. You should see the following:

Input Sequence or Mass (Da): ☒ b-ion ☐ y-ion ☐ non-specific

AM<OX M>

Peptide Mass: 219.08 Charge: 1 M/Z Mass: 219.08

<- 1 ->

To view additional PTMs from the unimod list, select the “unimod” box. You can delete a PTM using the “Delete PTM” button. You can also edit or create a new PTM using the “Edit PTM” or “New PTM” buttons, respectively.

You can also use the advanced options menu to add a water or proton to your sequence. The “Back” button will remove the last amino acid or modification that you have added to your sequence or click on the “Clear” button to start again.

19 About Bioinformatics Solutions Inc.

BSI provides advanced software tools for analysis of biological data.

Bioinformatics Solutions Inc. develops advanced algorithms based on innovative ideas and research, providing solutions to fundamental bioinformatics problems. This small, adaptable group is committed to serving the needs of pharmaceutical, biotechnological and academic scientists and to the progression of drug discovery research. The company, founded in 2000 in Waterloo, Canada, comprises a select group of talented, award-winning developers, scientists, computer and biological scientists as well as valuable consultants and beta-testers.

At BSI, groundbreaking research and customer focus go hand in hand on our journey towards excellent software solutions. We value an intellectual space that fosters learning and an understanding of current scientific knowledge. With an understanding of theory, we can focus our talents on providing solutions to difficult, otherwise unsolved problems that have resulted in research bottlenecks. At BSI, we are not satisfied with a solution that goes only partway to solving these problems; our solutions must offer something more than existing software.

The BSI team recognizes that real people will use our software tools. As such, we hold in principle that it is not enough to develop solely on theory; we must develop with customer needs in mind. We believe the only solution is one that incorporates quality and timely results, a satisfying product experience, customer support and two-way communication. So then, we value market research, development flexibility and company-wide collaboration, evolving our offerings to match the market/user's needs.

Efficient and concentrated research, development, customer focus and market analysis have produced: PEAKS software for protein and peptide identification from tandem mass spectrometry data, RAPTOR and PROSPECT Pro software for threading based 3D protein structure prediction and PatternHunter software for all types of homology search sequence comparison.

20 PEAKS Software License

This is the same agreement presented on installation. It is provided here for reference only.

If we are evaluating a time limited trial version of PEAKS and we wish to update the software to the full version, we must purchase PEAKS and obtain a full version registration key.

1. License. Subject to the terms and conditions of this Agreement, Bioinformatics Solutions (BSI) grants to you (Licensee) a non-exclusive, perpetual, non-transferable, personal license to install, execute and use one copy of PEAKS (Software) on one single CPU at any one time. Licensee may use the Software for its internal business purposes only.

2. Ownership. The Software is a proprietary product of BSI and is protected by copyright laws and international copyright treaties, as well as other intellectual property laws and treaties. BSI shall at all times own all right, title and interest in and to the Software, including all intellectual property rights therein. You shall not remove any copyright notice or other proprietary or restrictive notice or legend contained or included in the Software and you shall reproduce and copy all such information on all copies made hereunder, including such copies as may be necessary for archival or backup purposes.

3. Restrictions. Licensee may not use, reproduce, transmit, modify, adapt or translate the Software, in whole or in part, to others, except as otherwise permitted by this Agreement. Licensee may not reverse engineer, decompile, disassemble, or create derivative works based on the Software. Licensee may not use the Software in any manner whatsoever with the result that access to the Software may be obtained through the Internet including, without limitation, any web page. Licensee may not rent, lease, license, transfer, assign, sell or otherwise provide access to the Software, in whole or in part, on a temporary or permanent basis, except as otherwise permitted by this Agreement. Licensee may not alter, remove or cover proprietary notices in or on the Licensed Software, or storage media or use the Licensed Software in any unlawful manner whatsoever.

4. Limitation of Warranty. THE LICENSED SOFTWARE IS PROVIDED AS IS WITHOUT ANY WARRANTIES OR CONDITIONS OF ANY KIND, INCLUDING BUT NOT LIMITED TO WARRANTIES OR CONDITIONS OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. LICENSEE ASSUMES THE ENTIRE RISK AS TO THE RESULTS AND PERFORMANCE OF THE LICENSED SOFTWARE.

5. Limitation of Liability. IN NO EVENT WILL LICENSOR OR ITS SUPPLIERS BE LIABLE TO LICENSEE FOR ANY INDIRECT, INCIDENTAL, SPECIAL, OR CONSEQUENTIAL

DAMAGES WHATSOEVER, EVEN IF THE LICENSOR OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE OR CLAIM, OR IT IS FORESEEABLE. LICENSOR'S MAXIMUM AGGREGATE LIABILITY TO LICENSEE SHALL NOT EXCEED THE AMOUNT PAID BY LICENSEE FOR THE SOFTWARE. THE LIMITATIONS OF THIS SECTION SHALL APPLY WHETHER OR NOT THE ALLEGED BREACH OR DEFAULT IS A BREACH OF A FUNDAMENTAL CONDITION OR TERM.

6. Termination. This Agreement is effective until terminated. This Agreement will terminate immediately without notice if you fail to comply with any provision of this Agreement. Upon termination, you must destroy all copies of the Software. Provisions 2,5,6,7 and 10 shall survive any termination of this Agreement.

7. Export Controls. The Software is subject at all times to all applicable export control laws and regulations in force from time to time. You agree to comply strictly with all such laws and regulations and acknowledge that you have the responsibility to obtain all necessary licenses to export, re-export or import as may be required.

8. Assignment. Customer may assign Customer's rights under this Agreement to another party if the other party agrees to accept the terms of this Agreement, and Customer either transfer all copies of the Program and the Documentation, whether in printed or machine-readable form (including the original), to the other party, or Customer destroy any copies not transferred. Before such a transfer, Customer must deliver a hard copy of this Agreement to the recipient.

9. Maintenance and Support. BSI will provide technical support for a period of thirty (30) days from the date the Software is shipped to Licensee. Further maintenance and support is available to subscribers of BSI's Maintenance plan at BSI's then current rates. Technical support is available by phone, fax and email between the hours of 9 am and 5 pm, Eastern Time, excluding statutory holidays.

10. Governing Law. This Agreement shall be governed by and construed in accordance with the laws in force in the Province of Ontario and the laws of Canada applicable therein, without giving effect to conflict of law provisions and without giving effect to United Nations Convention on contracts for the International Sale of Goods.

For technical support issues not found in this manual, please contact either your Sales Representative or any of our support service resources:

email: support@bioinfor.com

tel: (519) 885-8288

fax: (519) 885-9075

online: www.bioinfor.com/peaks

